Helmut Diehl

MODELL UND METHODEN ZUR VORAUSRECHNUNG VON RINDERPROZESSEN

Saarbrücken 1970

MODELL UND METHODEN ZUR VORAUSRECHNUNG VON RINDERPROZESSEN

I n a u g u r a l d i s s e r t a t i o n

zur Erlangung des akademischen Grades
eines
Doktors der Wirtschaftswissenschaft
der
Rechts- und Wirtschaftswissenschaftlichen Fakultät der Universität des Saarlandes

(doctor rerum politicarum)

vorgelegt von :
Diplom-Mathematiker
Helmut Diehl

UNIV.
LEIDEN
DIBL.

Tag der Disputation : 12. Februar 1970
Dekan : Prof. Dr. H. Hax
Erstberichterstatter : Prof. Dr. G. Menges
Zweitberichterstatter : Prof. Dr. H. Schneeweiß

VORWORT

Die Anregung zur vorliegenden Arbeit geht zurück auf Diskussionen, welche ich mit Herrn Prof. S.L. Louwes[1] über das Problem der rationellen Gestaltung statistischer Aktivität im allgemeinen und insbesondere im Bereich der Agrarstatistik führte. Das wesentliche Ergebnis dieser Diskussionen war die Erkenntnis, dass eine befriedigende Lösung dieses Problems nur durch eine strenge Orientierung an objektiven Kriterien und durch die Anwendung geeigneter quantitativer Methoden zu erreichen ist.[2]

Eine Möglichkeit, diese Gedanken auf konkrete praktische Probleme anzuwenden, bot sich im Zusammenhang mit der Aufstellung eines agrarstatistischen Arbeitsprogramms für die Mitgliedsstaaten der Europäischen Gemeinschaften[3], in welchem Erhebungen und kurzfristige Prognosen auf dem tierischen und pflanzlichen Sektor geplant bzw. reorganisiert und harmonisiert werden sollen. Im Rahmen der vorbereitenden Arbeiten zu diesem Programm wurde mir der Auftrag zuteil, das Problem der kurzfristigen Prognose auf dem Rindersektor zu untersuchen.
Es erschien mir zweckmässig, dieses Problem zunächst einmal aus theoretischer Sicht zu behandeln und dabei einige wichtige theoretisch-statistische Aspekte, darunter insbesondere auch den der Rationalisierung des statistischen Aufwandes für eine Prognose, sehr eingehend zu studieren. Dies geschieht in den Ausführungen dieser Arbeit.

Zu Dank verpflichtet bin ich Herrn Prof. S.L. Louwes für die mir im Verlauf der Bearbeitung in zahlreichen Gesprächen vermittelten Anregungen sowie den Herren Prof.Dr.G.Menges und Prof.Dr.H.Schneeweiss für wertvolle Hinweise in Bezug auf das theoretische Konzept.

---

1) Direktor der Agrarstatistik im Statistischen Amt der Europäischen Gemeinschaften.

2) Mit solchen Gedanken beschäftigen sich die Arbeiten [12] und [13] sowie die in gewissem Sinn als Ergänzungen anzusehenden Arbeiten [6] und [14] .

3) Dieses Arbeitsprogramm befindet sich in der Direktion Agrarstatistik des Statistischen Amtes der Europäischen Gemeinschaften in Vorbereitung.

INHALTSVERZEICHNIS

		Seite
§ 1	Einleitung	1
§ 2	Der Rinderprozess	5
	2.1 Das Phänomen	5
	2.2 Die mathematische Definition des Rinderprozesses	6
	2.3 Die Beobachtung und Darstellung von Rinderprozessen	8
	2.4 Die Bestimmungsfaktoren des Rinderprozesses und seine Fortschreibung	10
	2.5 Dem Rinderprozess ähnliche Prozesse	12
§ 3	Ein deterministisches Modell zur Vorausrechnung eines Rinderprozesses	13
	3.1 Das Vorausrechnungsmodell der Erwartungswerte	13
	3.2 Die im Modell implizierten Voraussetzungen	24
	3.3 Verfügbarkeit bzw. Bestimmbarkeit der Modelldaten	28
	3.4 Wert und Grenzen dieses deterministischen Modells	32
§ 4	Ein stochastisches Modell zur Vorausrechnung eines Rinderprozesses	35
	4.1 Das Modell	35
	4.2 Die Konstruktion von Wahrscheinlichkeitsverteilungen φ_i für die erklärenden Grössen des Modells	42
§ 5	Der entscheidungstheoretische Aspekt bei der Vorausrechnung von R_{lat_p}	48
	5.1 Vorbemerkung	48
	5.2 Die Prognose von R_{lat_p} als Entscheidungsproblem - die Bayessche Prognose und das Bayessche Risiko	49

		Seite
5.3	Die optimale Prognose für R_{lat_p} bei speziellen Verlustfunktionen	52
5.4	Die praktische Relevanz des Begriffs der Verlustfunktion beim Problem der Vorausrechnung von Rinderprozessen	58
5.5	Die den weiteren Untersuchungen zugrunde gelegte Verlustfunktion	60
§ 6	Bestimmung des Mittelwertes, der Streuung und des Variationskoeffizienten von R_{lat_p}	61
§ 7	Die Optimierung von Informationsprogrammen zur Vorausrechnung von R_{lat_p}	68
7.1	Vorbemerkung	68
7.2	Das Problem der Optimierung von Informationsprogrammen im allgemeinen	69
7.3	Die allgemeine Lösung des Optimierungsproblems von Informationsprogrammen	70
7.4	Die Optimierung von Informationsprogrammen zur Vorausrechnung von R_{lat_p}	75
§ 8	Zusammenfassung	84
Literaturverzeichnis		88

§ 1 EINLEITUNG

Für die Analyse und Steuerung des Rindermarktes ist es wichtig, ausser über laufende Bestandsstatistiken auch über laufende Prognosen zu verfügen.

Solche Prognosen können auf kurzen, mittleren und langen Termin erfolgen. Lang- und mittelfristige Prognosen implizieren, durch die Prognosezeitspanne bedingt, ausser gewissen biologischen und zootechnischen[1] Faktoren der Rinderhaltung insbesondere auch ökonomische Elemente des Rindermarktes wie Preise und Elastizitäten von Rindfleisch und Milch und schliesslich auch die Einflüsse agrarpolitischer Massnahmen. Solche Prognosen erfordern die Aufstellung geeigneter ökonometrischer Modelle.[2] Kurzfristige Prognosen dagegen beziehen sich auf eine Zeitspanne, welche so bemessen ist, dass Reaktionen auf das Marktgeschehen und die Agrarpolitik vermutlich noch keinen oder nur unbedeutenden Einfluss ausüben, so dass nur die relativ konstanten biologischen und zootechnischen Faktoren des Phänomens von Bedeutung sind. Die kurzfristige Prognose stellt also eine sich nur an diesen Faktoren orientierende Vorausrechnung der Eigenentwicklung eines Bestandes dar und erfordert also statt eines ökonometrischen Modells ein biotechnisches Modell. Zur Unterscheidung von der durch das ökonomische Geschehen mitbestimmten mittel- und langfristigen Prognose und zur Betonung ihres mehr technisch statistischen Charakters soll die kurzfristige Prognose als Vorausrechnung bezeichnet werden.

Gegenstand der vorliegenden Arbeit ist eine Untersuchung der Problematik der im Sinne der obigen Definition zu verstehenden Vorausrechnung von Rinderbeständen unter besonderer Beachtung einiger wichtiger theoretisch-statistischer Aspekte.

---

1) D.h. mit der Tierhaltung zusammenhängend.

2) Zur Theorie und Praxis der Erstellung von Prognosen an Hand ökonometrischer Modelle vgl. z.B. [16], [17], [20].

Von einer allgemeinen Betrachtung über den sogenannten Rinderprozess ausgehend, wird erst einmal festgestellt werden, welche Faktoren einen Rinderprozess bestimmen und wie bei Bekanntsein derselben ein Rinderbestand fortgeschrieben werden kann. Für die Fortschreibung eines Bestandes in die Zukunft oder, wie wir sagen, Vorausrechnung des Prozesses wird durch entsprechende Modifikation des Fortschreibungsverfahrens zunächst ein für bescheidene Zwecke geeignetes deterministisches Modell entwickelt, welches die stochastischen Aspekte des Phänomens ausser Acht lässt. Eine Erörterung der diesem Modell unterstellten Voraussetzungen sowie eine kurze Prüfung der praktischen Zugänglichkeit der Bestimmungsgrössen des Modells sollen zeigen, dass das Modell nicht praxisfremd ist.

Ein solches deterministisches Modell kann indes nur bescheidene Ansprüche befriedigen und lässt zudem durch die Vernachlässigung stochastischer Momente bedeutende Aspekte verloren gehen. Für die weiteren Untersuchungen wird deshalb, dem stochastischen Charakter des Phänomens Rechnung tragend, ein stochastisches Modell zur Vorausrechnung eines Rinderprozesses entwickelt. Erst auf der Grundlage dieses Modells eröffnen sich für die Untersuchung wichtiger mit der Vorausrechnung zusammenhängender Probleme die günstigen Perspektiven und für ihre Lösung die geeigneten Wege. Insbesondere ist dabei an die folgenden im Rahmen dieser Arbeit zu behandelnden Probleme gedacht.

Im Zusammenhang mit der stochastischen Vorausrechnung eines Rinderprozesses ist es zunächst wichtig, aus gegebenen Informationen eine Wahrscheinlichkeitsverteilung für die jeweils vorauszurechnende Grösse zu konstruieren - nämlich die sogenannte Prognoseverteilung . Das entwickelte stochastische Modell zeigt, wie sich dieselbe aus den Wahrscheinlichkeitsverteilungen der selbst wieder stochastischen erklärenden Grössen zusammensetzt. Für die Konstruktion dieser Wahrscheinlichkeitsverteilungen aus gegebenen Informationen ist schliesslich die statistische Inferenz zuständig.

Wenn eine Prognoseverteilung zur Verfügung steht, dann stellt sich die Frage, welche in der Angabe eines einzigen Wertes bestehende Prognose angesichts bestimmter Konsequenzen bei Fehlprognosen die optimale ist. Die Antwort auf diese Frage ist im Rahmen eines entscheidungstheoretischen Ansatzes zu suchen : bei Vorgabe einer Verlustfunktion und bekannter Prognoseverteilung kann mit Hilfe des Bayesschen Entscheidungskriteriums die optimale sogenannte Bayessche Prognose bestimmt werden.

Ein wichtiges Indiz zur Beurteilung des Ausmasses der Ungewissheit in Verbindung mit den Konsequenzen bei Fehlprognosen liefert das sogenannte Bayessche Risiko, welches sich mit der Bayesschen Prognose verbindet (ein Spezialfall des Bayesschen Risikos ist als mittlerer quadratischer Fehler bekannt). Bayessche Prognose und Bayessches Risiko sind aus dieser entscheidungstheoretischen Perspektive die einzigen im Zusammenhang mit einem Prognoseproblem interessierenden Grössen. Für die Prognoseverteilung bedeutet dies, dass man über sie nur soviel zu wissen braucht, dass man diese beiden Grössen angeben kann.

Für weitere Betrachtungen ist es erforderlich, eine geeignete Verlustfunktion zu spezifizieren. Auf der Basis derselben sind dann die Bayessche Prognose und das Bayessche Risiko als Funktionen spezieller Parameter der Prognoseverteilung zu ermitteln. Diese Parameter sind schliesslich an Hand des stochastischen Modells als Funktionen entsprechender Parameter der Verteilungen der erklärenden Grössen darzustellen.

Schliesslich ist im Zusammenhang mit dem Problem der Vorausrechnung von Rinderprozessen die folgende Aufgabenstellung von grosser Bedeutung.

Das Ausmaß der Ungewißheit, mit welcher ein Vorausrechnungswert behaftet ist, hängt von dem jeweils zugrunde liegenden sogenannten Informationsprogramm ab. In der Tatsache, dass ein Informationsprogramm die zur Einschränkung von Ungewißheit erforderlichen Informationen liefern kann, liegt ein gewisser Nutzen oder Wert desselben. Diesem Nutzen eines Informationsprogramms stehen andererseits aber auch die sich mit ihm verbindenden Kosten gegenüber und es stellt sich schliesslich das Problem, dasjenige Informationsprogramm zu finden, für welches der Nutzeffekt und die Kosten in einem optimalen Verhältnis zueinander stehen. In etwas modifizierter Form kann dieses Problem auch darin bestehen, zu einem vorgegebenen Budget das Informationsprogramm mit dem grössten Nutzeffekt ausfindig zu machen. Zur Lösung solcher Optimierungsprobleme ist ebenfalls von entscheidungstheoretischen Hilfsmitteln Gebrauch zu machen.

Bei diesem letzten Problem der Optimierung von Informationsprogrammen handelt es sich ganz allgemein gesprochen um ein Problem der Rationalisierung statistischer Aktivität. Diese Rationalisierung wird in diesem besonderen Fall an dem speziellen Verwendungszweck der Ergebnisse der Aktivität für eine Vorausrechnung orientiert. Die Diskussion um dieses Problem soll letztlich nochmals deutlich machen, wie wichtig es für die rationelle Gestaltung statistischer Tätigkeit ist, sich des Ziels dieser Tätigkeit bewusst zu sein und aus dieser Perspektive schliesslich nach den erforderlichen Optimierungskriterien zu suchen.

§ 2 DER RINDERPROZESS

2.1 Das Phänomen

Es ist zweckmässig, den folgenden Untersuchungen eine kurze phänomenologische Betrachtung über das Rind vorauszuschicken. Das Rind liefert zwei für die Ernährung wichtige Produkte : Milch und Fleisch, wobei je nach dem primären Zweck der Rinderhaltung das eine als willkommenes Nebenprodukt des anderen aufgefasst werden kann. Bei der Rinderhaltung liegt also ein 2-Produkt-System vor - ein Faktum, welches bei gewissen Analysen und langfristigen Prognosen des Bestandes zu berücksichtigen wäre, aber für unser Vorhaben der kurzfristigen Vorausrechnung des Bestandes ohne Bedeutung ist. Ohne wesentliche Aspekte zu verlieren, können wir unsere Betrachtungen auf das Rind als Fleischproduzenten beschränken, denn jedes Rind, ganz gleich ob es a priori für die Fleischproduktion bestimmt ist oder nicht, wird (wenn es nicht eines natürlichen Todes stirbt) eines Tages zur Fleischerzeugung beitragen.

Wir betrachten zunächst einmal den Lebensweg eines einzelnen Rindes. Derselbe kann, je nach dem Ermessen des Viehhalters und je nach Gegebenheit gewisser biologischer Voraussetzungen (z.B. ob es sich um ein männliches oder weibliches Tier handelt) einen sehr verschiedenartigen Verlauf nehmen. Nach 9 Monaten Tragzeit fast immer als Einzeltier zur Welt gebracht, kann das Tier entweder bereits als junges Kalb oder etwas später als Mastkalb geschlachtet werden, oder es kann zunächst als Kuh für die Regeneration und Milcherzeugung bestimmt werden, bevor es in höherem Alter geschlachtet wird; schliesslich kann es auch als Zuchtstier gehalten oder als Masttier ausschliesslich zum Zwecke eines hohen Fleischertrages gemästet werden.

Solche zootechnischen Formen des Lebensweges eines Rindes bezeichnen wir als Laufbahnen. Dieser Begriff der Laufbahn eines Rindes und sein Alter zum Zeitpunkt der Schlachtung, das sogenannte Schlachtalter, werden für die Belange unserer folgenden eine Gesamtheit von Rindern betreffenden globalen Betrachtungen die einzig erheblichen Merkmale des Lebensweges eines Rindes sein.

Objekt unserer folgenden Untersuchungen wird also, wie gesagt, nicht der individuelle Lebensweg eines Tieres sein, sondern die Summe der Lebenswege aller Rinder innerhalb eines bestimmten Haltungsgebietes und eines bestimmten Zeitabschnittes - eine solche Summe wollen wir als Rinderprozess bezeichnen.

2.2 Die mathematische Definition des Rinderprozesses

Von der obigen Summendefinition des Rinderprozesses abweichend, wollen wir jetzt eine unseren Zwecken angepasstere Definition des Rinderprozesses geben :

> Ein Rinderprozess ist gegeben durch die zeitliche Entwicklung des nach gewissen vorgegebenen Kriterien gegliederten momentanen Rinderbestandes innerhalb eines bestimmten Haltungsgebietes.

Als Gliederungskriterien des Bestandes legen wir das Alter eines Tieres und seine Laufbahnzugehörigkeit fest.

Bezüglich des Laufbahnbegriffs wollen wir noch folgendes vereinbaren : Eine Laufbahn kann nach gewissen Kriterien wie Rasse, Region, Betriebsstruktur oder irgendwelchen anderen nicht das Alter implizierende Kriterien in sogenannte Unterlaufbahnen zerfallen. Alle Laufbahnen und Unterlaufbahnen haben jedoch mit der Altersklasse 0 zu beginnen. Wenn eine

Laufbahn nach dem Alter zerlegt wird, dann bezeichnen wir die entstehenden Teilmengen als Laufbahnaltersgruppen, welche jedoch keine eigenständige Laufbahn mehr darstellen. So bildet z.B. die Gesamtheit der Milchkühe lediglich eine Laufbahnaltersgruppe. Um diese Laufbahnaltersgruppe zu einer Laufbahn zu ergänzen, hat man noch diejenigen Tiere mit einzubeziehen, welche eines Tages Milchkühe sein werden.

Zum Zwecke einer mathematischen Formulierung der Begriffe Rinderbestand und Rinderprozess führen wir die folgenden Bezeichnungen ein :

Das Merkmal Laufbahn bezeichnen wir mit l, die Menge aller Laufbahnen mit L. Das Merkmal Alter bezeichnen wir mit a, die Gesamtheit aller Altersstufen, bzw. die mögliche Lebensdauer, mit A. Der Relevanzzeitpunkt wird mit t, der Relevanzzeitraum mit T bezeichnet. Mit R_{lat} bezeichnen wir die Anzahl der Tiere, welche zum Zeitpunkt t der Laufbahn l angehören und das Alter a haben. Die innerhalb eines bestimmten Gebietes zum Zeitpunkt t über die Laufbahnen und Altersklassen genommene Gesamtheit der Grössen R_{lat} bezeichnen wir als Rinderbestand in t. Wir symbolisieren den Rinderbestand zum Zeitpunkt t mit $R_t = \{R_{lat};\ (l,a) \in (L \times A)\}$. Der Rinderprozess ist schliesslich die über den Zeitraum T genommene Gesamtheit der Rinderbestände R_t. Wir symbolisieren den Rinderprozess mit $R = \{R_{lat};\ (l,a,t) \in (L \times A \times T)\}$.

Ein Rinderprozess ist also in unserem Sinne eindeutig beschrieben, wenn für jeden Zeitpunkt seines Relevanzzeitraumes bekannt ist, wieviele Tiere sich in jeder Altersklasse einer jeden Laufbahn befinden. In der Praxis sind natürlich solche zeitlich kontinuierlichen Angaben nicht möglich, man hat sich auf die Registrierung des Bestandes in diskreten meist äquidistant liegenden Zeitpunkten t_o, t_1, t_2, ... zu beschränken.

An dieser Stelle soll schon erwähnt werden, dass ein Rinderprozess als ein stochastischer Prozess aufzufassen ist, da seine Entwicklung durch stochastische Elemente mitbestimmt wird. Im Gegensatz zu dieser Auffassung würde diejenige stehen, dass es sich beim Rinderprozess um einen deterministischen Prozess handelt, bei welchem sich die Entwicklung streng nach mathematischen Funktionen vollzieht.[1] Diese Auffassung ist zwar dem Phänomen nicht angepasst, kann aber, wie wir später sehen werden (vgl. § 3.4), unter gewissen Umständen auch eine Berechtigung haben.

Aus der Sicht der Praxis sind im Zusammenhang mit dem Phänomen des Rinderprozesses 3 Probleme von besonderem Interesse :

1) Die Beobachtung und Darstellung eines Rinderprozesses.
2) Die Frage, welche Faktoren die Entwicklung eines Rinderprozesses bestimmen und wie derselbe fortgeschrieben werden kann.
3) Die Frage, wie ein Rinderprozess vorausberechnet werden kann.

Das erste Problem werden wir wegen seiner hier untergeordneten Bedeutung nur streifen. Die Beantwortung der Frage 2 ist wesentlich für die Behandlung des dritten Problems und wird etwas mehr Aufmerksamkeit erfordern. Die Frage 3 wird schliesslich den Ausgangspunkt der weiteren Untersuchungen darstellen.

2.3 Die Beobachtung und Darstellung eines Rinderprozesses

Die Grundlage für Analysen von Problemen der Rinderhaltung und des Rindermarktes sowie für die Vorbereitung von Entscheidungen auf diesem Sektor sind die aus laufenden Bestands-

1) Wegen näherer Erklärungen zu stochastischen und deterministischen Prozessen vgl.[1],[8],[33].

statistiken hervorgehenden Informationen. Die Führung einer laufenden Bestandsstatistik kennt zwei Probleme : die laufende Beobachtung oder Erhebung der Bestände und die Darstellung derselben. Nach unserem Konzept wären die Bestandsgrössen R_{lat} in den verschiedenen Laufbahnen und Altersklassen zu bestimmten Zeitpunkten Gegenstand der Beobachtung. Wir wollen hier nicht auf die zahlreichen erhebungstechnischen Probleme eingehen, sondern nur erwähnen, dass die Angaben über R_{lat} im Prinzip durch zwei Verfahren zustande kommen können : entweder durch Vollzählungen oder durch Stichprobenverfahren mit anschliessender Hochrechnung.[1)]

Die Darstellung der erhobenen Bestände R_{lat} erfolgt zweckmässigerweise in der Form, dass man für jeden Erhebungszeitpunkt die gewonnenen Angaben in zweidimensionale Tabellen folgender Art einträgt :

Bestand zum Zeitpunkt t

Alter \ Laufbahn	l_1	l_2	. . .	l_k
a_1	$R_{l_1 a_1 t}$	$R_{l_2 a_1 t}$	. . .	$R_{l_k a_1 t}$
a_2	$R_{l_1 a_2 t}$	$R_{l_2 a_2 t}$	. . .	$R_{l_k a_2 t}$
.				
.				
.				
a_r	$R_{l_1 a_r t}$	$R_{l_2 a_r t}$	. . .	$R_{l_k a_r t}$

Eine Folge solcher Tabellen über einen gewissen Zeitraum und ein gewisses Gebiet gibt eine Beschreibung des betreffenden Rinderprozesses.

1) Vgl. hierzu [3], [30].

2.4 Die Bestimmungsfaktoren des Rinderprozesses und seine Fortschreibung

Bei der Analyse eines Rinderprozesses stellt sich als erstes die Frage, welche Faktoren die Veränderungen eines Bestandes in der Zeit bewirken.

Bei der Untersuchung dieses Problems beschränken wir unsere Betrachtungen auf einen nach aussen abgeschlossenen Prozess, welcher Zugänge nur durch eigene Regeneration und Abgänge nur durch Schlachtungen erfährt. Insbesondere sind damit Importe und Exporte in Bezug auf das betrachtete Haltungsgebiet ausgeschlossen. Ausserdem impliziert dies noch die Annahme, dass keine natürliche Mortalität vorliegt. An späterer Stelle werden wir ausführlich auf die hier getroffenen und die im folgenden noch zu unterstellenden Voraussetzungen und Einschränkungen zu sprechen kommen.

Die Entwicklung eines Rinderprozesses ist das Ergebnis einer an einen bestimmten Anfangsbestand sich anschliessenden laufenden Bilanz zwischen Neuzugängen und Schlachtungen, wobei allerdings bei dieser Bilanzierung nach Laufbahnen, und was die Schlachtungen betrifft, auch nach Altersklassen zu differenzieren ist (die Neuzugänge treten automatisch in die Anfangsaltersklasse ein).

Der Verlauf eines Rinderprozesses ist also bestimmt durch :

1. Einen Anfangsbestand
2. Die Anzahl der Neuzugänge in jedem Zeitpunkt in den einzelnen Laufbahnen
3. Die Anzahl der Schlachtungen in jedem Zeitpunkt in den einzelnen Laufbahnen und Altersklassen.

Wenn über diese drei Bestimmungsstücke die entsprechenden Angaben vorliegen, dann kann ein Rinderbestand von einem Ausgangspunkt t_o nach einem späteren Zeitpunkt t_1 fortgeschrieben

werden. Zum Zwecke einer mathematischen Formulierung einer solchen Fortschreibung führen wir die folgenden Bezeichnungen ein: Die Anzahl der Neuzugänge in der Laufbahn l zum Zeitpunkt t bezeichnen wir mit N_{lt} [1] und die Anzahl der Schlachtungen in der Altersklasse a der Laufbahn l zum Zeitpunkt t mit S_{lat}.

Für die Berechnung des Bestandes $\{R_{lat_1}\}$ aus dem Bestand $\{R_{lat_o}\}$ an Hand von Angaben über die Neuzugänge N_{lt} und die Schlachtungen S_{lat} gilt dann die folgende Fortschreibungsformel [2)3)]:

$$(1)\quad R_{lat_1} = \begin{cases} R_{l,a-t_1+t_o,t_o} - \int_{t_o}^{t_1} S_{l,a-t_1+t,t}\,dt & \text{für } a \geq t_1 - t_o \\ N_{l,t_1-a} - \int_{t_1-a}^{t_1} S_{l,a-t_1+t,t}\,dt & \text{für } a < t_1 - t_o. \end{cases}$$

Diese Fortschreibungsformel dient dazu, ohne weitere Bestandserhebungen nur an Hand von Zugangs- und Schlachtungsabgaben die Entwicklung eines Rinderprozesses zu verfolgen.

Eine Fortschreibung an Hand der Formel (1) kann natürlich nur von einem Zeitpunkt der Vergangenheit bis höchstens in die

---

1) Neuzugänge sind Tiere des Alters o, wir können deshalb auch schreiben $N_{lt} = R_{l,o,t}$.

2) Wegen einer allgemeinen mengentheoretischen Formulierung des Fortschreibungsproblems vgl. [18], Seite 67 ff.

3) Der Einfachheit der Darstellung wegen werden im folgenden ausschliesslich kontinuierliche Zeit- und Altersparameter verwendet, nur der Laufbahnparameter l ist eine diskrete Variable. Die Grössen N_{lt}, S_{lat} und R_{lat} sind dann als Differentiale aufzufassen: N_{lt} ist die Anzahl der Neuzugänge pro infinitesimaler Zeiteinheit zum Zeitpunkt t in der Laufbahn l; S_{lat} ist die Anzahl der Schlachtungen pro infinitesimaler Zeiteinheit zum Zeitpunkt t und pro infinitesimaler Alterseinheit im Alter a in der Laufbahn l; R_{lat} ist die zum Zeitpunkt t in der Laufbahn l vorhandene Anzahl von Tieren pro infinitesimaler Alterseinheit im Alter a.

Gegenwart erfolgen, da sie das Vorliegen von Angaben über Neuzugänge und Schlachtungen voraussetzt. Eine exakte Fortschreibung auf der Basis wahrer Werte über Neuzugange und Schlachtungen ist über den Gegenwartszeitpunkt hinaus nicht möglich. In den weiteren Untersuchungen werden wir jedoch sehen, wie man durch geschickte Substitution gewisser Elemente der obigen Fortschreibungsformel zu einem Modell gelangt, welches es ermöglicht, einen Rinderprozeß auch in die Zukunft fortzuschreiben oder -wie wir sagen werden- vorauszuberechnen.

2.5 Dem Rinderprozeß ähnliche Prozesse

Bevor wir fortfahren, wollen wir noch feststellen, dass unsere Betrachtungen keineswegs auf den Rinderprozess allein beschränkt bleiben müssen, sondern ohne weiteres verallgemeinert und auf zahlreiche andere Prozesse übertragen werden können. Im Prinzip können die hier angestellten Untersuchungen auf alle Prozesse angewendet werden, welche man als sogenannte Geburts- Todes- Prozesse[1)] deuten kann, d.h. als Prozesse, die in ihrer Entwicklung durch laufende Zugänge und laufende durch Altern sich anbahnende Abgänge bestimmt sind. Solche Prozesse finden sich in fast allen Gebieten der Technik, Natur, Wirtschaft und Gesellschaft.

Insbesondere auf dem Agrarsektor gibt es noch eine Anzahl dem Rinderprozess sehr ähnlicher Prozesse. In erster Linie ist dabei an die Entwicklung von Schweinebeständen zu denken. Bei der Uebertragung unserer Gedankengänge auf den Schweineprozess wird allerdings sofort auffallen, dass die einzelnen Begriffe und deren Bedeutung eine gewisse Verschiebung erfahren. Zum Beispiel ist beim Rinderprozeß die Anzahl der von einer Kuh pro Jahr geborenen Kälber nahezu eine Konstante (nämlich 1), beim Schweineprozeß dagegen ist die Anzahl der von einer Sau jährlich geborenen Ferkel (bis 3 mal jährlich bis zu 15 Ferkel pro Wurf) eine Zufallsgröße mit erheblicher Streuung.

---

1) Vgl. hierzu [9], Seite 244 ff.

Umgekehrt schwankt das Schlachtalter bei Mast-Schweinen um 1-2 Monate während bei Rindern in diesem Punkt viel grössere Unsicherheit herrscht. Der beim Rinderprozess sehr wichtige Laufbahnbegriff ist bei Schweinen kaum von Bedeutung, da man lediglich Mutterschweine und Masttiere unterscheidet. Ein bedeutender Unterschied zwischen beiden besteht in der Regenerationsfähigkeit, welche bei Schweinen durch die sehr hohe Ferkelzahl pro Muttertier und Jahr ungleich höher ist als bei Rindern. Solche Unterschiede betreffen jedoch nicht die Logik des allgemeinen Modells solcher Prozesse, sondern nur die sich entsprechenden Details.

Die interessanteste Parallele zum Rinderprozess stellt die Bewegung von Bevölkerungsbeständen dar. Auf Grund jahrzehntelanger exaktester Statistiken liegen hier präzise Kenntnisse über alle Bestimmungsfaktoren vor, was der Entwicklung von Analyse- und Prognosemodellen sehr entgegenkommt.

Wir wollen keine Aufzählung weiterer Parallelen vornehmen, sondern es bei diesen kurzen Erwähnungen belassen, deren Absicht es lediglich war, auf die Verwandtschaft unseres Problems zu anderen aufmerksam zu machen.

§ 3 EIN DETERMINISTISCHES MODELL ZUR VORAUSRECHNUNG EINES RINDERPROZESSES

3.1 Das Vorausrechnungsmodell der Erwartungswerte

Oben hatten wir gesehen, dass für die Entwicklung eines Rinderprozesses ein Ausgangsbestand und die zeitlich darnach liegenden Neuzugänge und Schlachtungen in den Laufbahnen und Altersklassen massgeblich sind. Dieselben Faktoren sind natürlich auch für die zukünftige Entwicklung

eines Prozesses bestimmend, und der logische Inhalt der Fortschreibungsformel (1) bleibt im Prinzip auch gültig, nur müssen die fehlenden Angaben über wirkliche Werte der Neuzugänge und Schlachtungen durch Schätzungen der zukünftigen Werte ersetzt bzw. durch prozessendogene vorherbestimmte Grössen erklärt werden. Dies wird zu einer gewissen Modifikation der Formel (1) führen.

Zunächst wollen wir den Abschlachtungsfaktor betrachten :

Anstelle des Begriffs der Anzahl von Schlachtungen S_{lat} werden wir den Begriff der Schlachtaltersverteilung einführen. Wann ein Tier, das sich in einer gewissen Laufbahn befindet, geschlachtet wird, ist in das recht variable Ermessen des Viehhalters gestellt. Zahlreiche nicht erkennbare oder nicht messbare Faktoren beeinflussen diese Entscheidung, zumindest was das einzelne Tier betrifft. Das individuelle Schlachtalter, das wir mit α bezeichnen, kann also als eine stochastische Veränderliche gedeutet werden, von welcher man allerdings annehmen kann, dass sie einer gewissen Gesetzmässigkeit in Form einer Wahrscheinlichkeitsverteilung unterliegt. Diese Wahrscheinlichkeitsverteilung des Schlachtalters, welche wir mit $s_1(\alpha)$[1] bezeichnen, heisst die Schlachtaltersverteilung; mit dem Index 1 soll gekennzeichnet werden, dass diese Verteilung von der Laufbahn abhängt.

Aus $s_1(\alpha)$ lässt sich noch die sogenannte bedingte Schlachtaltersverteilung ableiten, welche die Schlachtaltersverteilung für ein Tier darstellt, welches bereits das Alter a erreicht hat. Diese bedingte Verteilung ist gegeben durch

$$s_1(\alpha | a) = s_1(\alpha) \Big/ \int_a^\infty s_1(\alpha)\, d\alpha, \alpha \geq a.$$

Ein weiterer wichtiger Begriff im Zusammenhang mit der

1) Wir fassen das Schlachtalter als kontinuierliche Grösse auf, $s_1(\alpha)$ ist dann genau genommen eine Wahrscheinlichkeitsdichtefunktion. Die Bezeichnung Wahrscheinlichkeitsverteilung (oder auch kurz : Verteilung) wollen wir jedoch als Oberbegriff für die diskrete und kontinuierliche Variante weiterhin verwenden.

Schlachtaltersverteilung ist die Lebenserwartung eines Tieres. Die absolute Lebenserwartung eines neugeborenen Tieres der Laufbahn 1 ist gegeben durch $E_1(\alpha) = \int_0^\infty \alpha \; s_1(\alpha) d\alpha$. Die bedingte Lebenserwartung eines Tieres der Laufbahn 1, welches bereits ein Alter a erreicht hat, ist gegeben durch $E_1(\alpha | a) = \int_a^\infty \alpha \; s_1(\alpha) \; d\alpha \Big/ \int_a^\infty s_1(\alpha) \; d\alpha$.

Wir geben jetzt noch eine empirische Definition der Schlachtaltersverteilung an Hand der zeitlichen Entwicklung eines Bestandes. Es sei R(a) ein Rinderbestand des Alters a, welcher mit zunehmendem Alter durch Abschlachtungen ständig abnimmt und schliesslich verschwindet. R(a) als Funktion von a (bzw. der Zeit) gesehen kennzeichnet die Bestands-Alters- Abhängigkeit. Wir definieren den Ausdruck

$$(2) \quad s(\alpha | a_o) \overset{def}{=} \left[- \frac{1}{R(a_o)} \cdot \frac{dR(a)}{da} \right]_{a=\alpha}, \quad \alpha \geqslant a_o,$$

als bedingte Schlachtaltersverteilung eines Tieres mit Alter a_o. Für $a_o = o$ erhält man die (nicht-bedingte) Schlachtaltersverteilung $s(\alpha) = \left[- \frac{1}{R(o)} \cdot \frac{dR(a)}{da} \right]_{a=\alpha}$. Umgekehrt kann man auch R(a) durch $s(\alpha | a)$ ausdrücken. Durch Integration von (2) erhält man

$$(2a) \quad R(a) = R(a_o) \left[1 - \int_{a_o}^{a} s \; (\alpha | a_o) \; d\alpha \right]$$

$$= R(a_o) \int_a^\infty s(\alpha | a_o) d\alpha, \quad (a \geqslant a_o).$$

Diese Definition der Schlachtaltersverteilung wird sich für die Praxis als sehr nützlich erweisen, insbesondere wegen der damit gegebenen Möglichkeit, die Schlachtaltersverteilung aus Prozessbeobachtungen zu ermitteln. Dazu ist es allerdings wichtig, über Prozessbeobachtungen in kurz hintereinander-

folgenden Zeitabständen zu verfügen.

Mit dem soeben eingeführten Begriff der Schlachtaltersverteilung, welche wir als vorherbestimmte Grösse betrachten, können wir die unbekannte Anzahl der Schlachtungen substituieren.

Nun zu den Neuzugängen N_{1t} :

Die Neuzugänge N_{1t} in den einzelnen Laufbahnen werden wir aus der Gesamtzahl der Kälbergeburten K_t [1] erklären. Wir stellen uns vor, dass die Zuweisung der Kälber in die einzelnen Laufbahnen zumindest gedanklich sofort nach der Geburt erfolgt und dass diese Zuweisung nach einer festen Wahrscheinlichkeitsverteilung γ_1, [2] der sogenannten Laufbahnverteilung, erfolgt. An die Stelle von N_{1t} wird der mathematische Erwartungswert von N_{1t} , nämlich $\gamma_1 \cdot K_t$, treten. Damit ist aber das Problem der Unkenntnis der zukünftigen Neuzugänge N_{1t} nur auf das der Unkenntnis der zukünftigen Kälbergeburten K_t zurückgeführt, und es bleibt immer noch die Frage offen, wie wir zu Kenntnissen über K_t für zukünftige Zeitpunkte gelangen. Da sich für die Ermittlung von K_t aus verfügbaren Grössen verschiedene Möglichkeiten anbieten, wollen wir dieses Problem als modellextern an späterer Stelle aufgreifen und zunächst die Zahl der Kälbergeburten K_t als Konstituenten in unser Modell aufnehmen.

Als Elemente, mit welchen wir ein Vorausrechnungsmodell für einen Rinderprozess aufstellen wollen, stehen uns also zur Verfügung :

1. Ein Ausgangsbestand zu einem Zeitpunkt t_o :

$$R_{t_o} = \{R_{lat_o}\} \quad .$$

2. Die Anzahl der Kälbergeburten zum Zeitpunkt t : K_t.

---

1) Bei kontinuierlichem Zeitparameter t stellt K_t die Anzahl der Kälbergeburten pro infinitesimaler Zeiteinheit zum Zeitpunkt t dar (vgl. hierzu Fußnote 3, Seite 11).

2) γ_1 gibt die Wahrscheinlichkeit für die Zuweisung eines neugeborenen Kalbes in die Laufbahn 1 an.

3. Die Laufbahnverteilung γ_l.

4. Die Schlachtaltersverteilungen in den Laufbahnen : $s_l(\alpha)$.

Das Problem ist folgendes :

Es soll vom Basiszeitpunkt t_o bei bekanntem Bestand R_{t_o} und bekannten Stücken K_t , γ_l , $s_l(\alpha)$ der Bestand zum Zeitpunkt t_p vorausberechnet werden.

Vorweg müssen wir noch klarstellen, was wir unter Vorausrechnung verstehen wollen :

Der zukünftige Bestand $R_{t_p} = \{R_{lat_p}\}$ ist zwar, allein schon durch den stochastischen Charakter des Schlachtalters und die stochastische Zuordnung der Kälbergeburten in die einzelnen Laufbahnen bedingt, eine stochastische Grösse. Wir beschränken uns jedoch zunächst darauf, den mathematischen Erwartungswert des Bestandes zum Zeitpunkt t_p zu ermitteln. Unter Vorausrechnung verstehen wir also die Berechnung der Erwartungswerte der Grössen R_{lat_p} bzw. der daraus abzuleitenden Grössen.

Wir berechnen als erstes den zu erwartenden Bestand zum Zeitpunkt t_p :

Zum Zeitpunkt t_p ist die Anzahl Tiere in der Laufbahn l mit einem Alter a, welches grösser oder gleich ist der Länge des Zeitraums (t_o,t_p), gleich der Anzahl derjenigen Tiere, welche in t_o das Alter $a - t_p + t_o$ haben und im Zeitraum (t_o, t_p) nicht geschlachtet werden. Der Erwartungswert dieser Anzahl ist gleich

$$R_{l,a-t_p+t_o,t_o} \cdot \int_a^\infty s_l(\alpha)\, d\alpha \Big/ \int_{a-t_p+t_o}^\infty s_l(\alpha)\, d\alpha \,.$$

Zum Zeitpunkt t_p ist die Anzahl Tiere in der Laufbahn 1 mit einem Alter a, welches kleiner ist als die Länge des Zeitraumes (t_o, t_p), gleich der Anzahl derjenigen Tiere, welche zum Zeitpunkt t_p - a zur Welt kommen, in die Laufbahn 1 eingeordnet und im Zeitraum (t_o, t_p) nicht geschlachtet worden. Der Erwartungswert dieser Anzahl ist gleich

$$\gamma_1 \cdot K_{t_p-a} \cdot \int_a^\infty s_1(\alpha)\, d\alpha \quad (a < t_p - t_o),$$

wobei man $\gamma_1 \cdot K_{t_p-a}$ auch mit $E(R_{1,o,t_p-a})$ bezeichnen kann.

Der zu erwartende Bestand in t_p ist also gegeben durch :

$$(3)\quad E(R_{lat_p}) = \begin{cases} R_{1,a-t_p+t_o} \cdot \int_a^\infty s_1(\alpha)\, d\alpha \Big/ \int_{a-t_p+t_o}^\infty s_1(\alpha)\, d\alpha & a \geq t_p - t_o \\ \gamma_1 \cdot K_{t_p-a} \cdot \int_a^\infty s_1(\alpha)\, d\alpha & a < t_p - t_o \end{cases}$$ [1]

---

1) Eine Analyse des Funktionsmechanismus dieses Modells dürfte von gewissem Interesse sein. Im Rahmen einer solchen Analyse wäre zu untersuchen, bei welcher Konstellation der Elemente des Modells der Prozess welche Entwicklung nimmt. Insbesondere wäre dabei herauszufinden, unter welchen Bedingungen der Prozess konstant und unter welchen er periodisch oder asymptotisch verläuft, wann er explodieren und wann er verschwinden würde. Des weiteren wäre es wichtig, die Wirkungsgeschwindigkeiten gewisser Veränderungen einzelner Elemente festzustellen.
Die Ergebnisse solcher Untersuchungen können von einigem praktischen Nutzen sein. Beispielsweise kann es im Zusammenhang mit der Planung bestimmter Maßnahmen auf dem Rindersektor von Bedeutung sein, auf Grund der Erfahrungen aus solchen Untersuchungen die Richtung und Geschwindigkeit der Wirkung von Maßnahmen schon im voraus kalkulieren zu können.

Zur Ermittlung der zu erwartenden Anzahl Tiere innerhalb der Laufbahnen bzw. insgesamt hat man in (3) entsprechend über a bzw. über a und l zu integrieren bzw. zu summieren.

Aus dem zu erwartenden Bestand (3) lässt sich die zu erwartende Anzahl Schlachtungen $E(S_{lat_p})$ zum Zeitpunkt t_p ableiten. Man erhält:

$$(4) \quad E(S_{lat_p}) = E(R_{lat_p})\, s_l(a) \Big/ \int_a^{\infty} s_l(\alpha)\, d\alpha ,$$

wobei je nach der Höhe des Alters für $E(R_{lat_p})$ die entsprechende Relation aus (3) gilt.

Wenn man übrigens in der Fortschreibungsformel (1) die Anzahl der Schlachtungen S_{lat} auf Grund von Formel (4) durch zu erwartende Schlachtungen ersetzt und den entstehenden Ausdruck unter Berücksichtigung der Relation (2a) sowie der Relation $N_{lt} = \gamma_l \cdot K_t$ entsprechend umformt, so gelangt man, wie intuitiv zu erwarten ist, zur obigen Formel (3). Wir werden dies ausführen.

Wenn man beachtet, dass dem Zeitpunkt t_1 der Formel (1) der Zeitpunkt t_p der Formel (3) entspricht, so gilt unter Vernachlässigung des Erwartungswertsymbols E folgendes:

Formel (1):

$$R_{lat_p} = \begin{cases} R_{1,a-t_p+t_o,t_o} - \int_{t_o}^{t_p} S_{1,a-t_p+t,t}\, dt & a \geq t_p - t_o \\ N_{1,t_p-a} - \int_{t_p-a}^{t_p} S_{1,a-t_p+t,t}\, dt & a < t_p - t_o \end{cases}$$

$$= \begin{cases} R_{1,a-t_p+t_o,t_o} - \int_{t_o}^{t_p} R_{1,a-t_p+t,t} \dfrac{s_1(a-t_p+t)}{\int_{a-t_p+t}^{\infty} s_1(\alpha)\,d\alpha}\,dt & a \geq t_p - t_o \\ N_{1,t_p-a} - \int_{t_p-a}^{t_p} R_{1,a-t_p+t,t} \dfrac{s_1(a-t_p+t)}{\int_{a-t_p+t}^{\infty} s_1(\alpha)\,d\alpha}\,dt & a < t_p - t_o \end{cases}$$

$$= \begin{cases} R_{1,a-t_p+t_o,t_o}\left[1 - \int_{t_o}^{t_p} \dfrac{R_{1,a-t_p+t,t}}{R_{1,a-t_p+t_o,t_o}} \cdot \dfrac{s_1(a-t_p+t)}{\int_{a-t_p+t}^{\infty} s_1(\alpha)\,d\alpha}\,dt\right] & a \geq t_p - t_o \\ \gamma_1 \cdot K_{t_p-a}\left[1 - \int_{t_p-a}^{t_p} \dfrac{R_{1,a-t_p+t,t}}{R_{1,o,t_p-a}} \cdot \dfrac{s_1(a-t_p+t)}{\int_{a-t_p+t}^{\infty} s_1(\alpha)\,d\alpha}\,dt\right] & a < t_p - t_o \end{cases}$$

(da $E(R_{1,o,t_p-a}) = \gamma_1 \cdot K_{t_p-a}$)

$$= \begin{cases} R_{1,a-t_p+t_o,t_o}\left[1 - \int_{t_o}^{t_p} \int_{a-t_p+t}^{\infty} s_1(\alpha \mid a-t_p+t_o)\,d\alpha \dfrac{s_1(a-t_p+t)}{\int_{a-t_p+t}^{\infty} s_1(\alpha)\,d\alpha}\,dt\right] & a \geq t_p - t_o \\ \gamma_1 \cdot K_{t_p-a}\left[1 - \int_{t_p-a}^{t_p} \int_{a-t_p+t}^{\infty} s_1(\alpha)\,d\alpha \dfrac{s_1(a-t_p+t)}{\int_{a-t_p+t}^{\infty} s_1(\alpha)\,d\alpha}\,dt\right] & a < t_p - t_o \end{cases}$$

(wegen Relation (2a))

$$= \begin{cases} R_{1,a-t_p+t_o,t_o}\left[1 - \int_{t_o}^{t_p} \frac{\int_{a-t_p+t}^{\infty} s_1(\alpha)d\alpha}{\int_{a-t_p+t_o}^{\infty} s_1(\alpha)d\alpha} \cdot \frac{s_1(a-t_p+t)}{\int_{a-t_p+t}^{\infty} s_1(\alpha)d\alpha}\, dt\right] & a \geq t_p - t_o \\ \gamma_1 \cdot K_{t_p-a}\left[1 - \int_{t_p-a}^{t_p} s_1(a-t_p+t)\, dt\right] & a < t_p - t_o \end{cases}$$

$$\left(\text{wegen } s_1(\alpha \mid a-t_p+t_o) = \frac{s_1(\alpha)}{\int_{a-t_p+t_o}^{\infty} s_1(\alpha)d\alpha}\right)$$

$$= \begin{cases} R_{1,a-t_p+t_o,t_o}\left[1 - \frac{\int_{t_o}^{t_p} s_1(a-t_p+t)\, dt}{\int_{a-t_p+t_o}^{\infty} s_1(\alpha)\, d\alpha}\right] & a \geq t_p - t_o \\ \gamma_1 \cdot K_{t_p-a}\left[1 - \int_{t_p-a}^{t_p} s_1(a-t_p+t)\, dt\right] & a < t_p - t_o \end{cases}$$

$$= \begin{cases} R_{1,a-t_p+t_o,t_o}\left[1 - \frac{\int_{a-t_p+t_o}^{a} s_1(\alpha)\, d\alpha}{\int_{a-t_p+t_o}^{\infty} s_1(\alpha)\, d\alpha}\right] & a \geq t_p - t_o \\ \gamma_1 \cdot K_{t_p-a}\left[1 - \int_{0}^{a} s_1(\alpha)\, d\alpha\right] & a < t_p - t_o \end{cases}$$

(wegen der Substitution $\alpha = a-t_p+t$)

$$= \begin{cases} R_{l,a-t_p+t_o,t_o} \cdot \dfrac{\int_a^\infty s_l(\alpha)\, d\alpha}{\int_{a-t_p+t_o}^\infty s_l(\alpha)\, d\alpha} & a > t_p - t_o \\ \gamma_l \cdot K_{t_p-a} \cdot \int_a^\infty s_l(\alpha)\, d\alpha & a < t_p - t_o . \end{cases}$$

Der letzte Ausdruck ist identisch mit Formel (3). Diese Identität demonstriert die Konsistenz des Inhalts der Formel (1) mit dem der Formel (3).

Wir wollen jetzt fortfahren mit einer weiteren Ableitung aus Formel (4). Wir ermitteln aus ihr die in t_p zu erwartende Gesamtanzahl der Schlachtungen $E(S_{t_p})$. Es ergibt sich:

$$(5) \quad E(S_{t_p}) = \sum_l \int_A E(S_{la t_p})\, da$$

$$= \sum_l \int_{t_p-t_o}^\infty R_{l,a-t_p+t_o,t_o} \cdot \frac{s_l(a)}{\int_{a-t_p+t_o}^\infty s_l(\alpha)\, d\alpha}\, da$$

$$+ \sum_l \int_o^{t_p-t_o} \gamma_l \cdot K_{t_p-a} \cdot s_l(a)\, da .$$

Will man auch noch das zu erwartende Gesamtschlachtgewicht $E(G_{t_p})$ berechnen, so hat man in Formel (5) die Anzahl Tiere mit laufbahn- und altersabhängigen mittleren Schlachtgewichten g_{la} zu multiplizieren und erhält:

$$(6)\quad E(G_{t_p}) = \sum_l \int_{t_p-t_o}^{\infty} g_{la}\, R_{l,a-t_p+t_o,t_o} \cdot \frac{s_l(a)}{\int_{a-t_p+t_o}^{\infty} s_l(\alpha)\,d\alpha}\, da$$

$$+ \sum_l \int_{o}^{t_p-t_o} g_{la} \cdot \gamma_l \cdot K_{t_p-a} \cdot s_l(a)\, da.$$

Zum Abschluss dieser Entwicklungen sei noch angedeutet, dass man an Hand des in Formel (3) vorausberechneten Bestandes bei Spezifikation der Laufbahnaltersgruppe Milchkühe (l_M) und gegebenen altersabhängigen mittleren Milcherträgen pro Tier m_a[1] den in t_p zu erwartenden Milchertrag $E(M_{t_p})$ vorausberechnen kann. Derselbe ist dann gegeben durch:

$$(7)\quad E(M_{t_p}) = \int_A m_a\, E(R_{l_M,a,t_p})\, da\ ,$$[2]

wobei man für $E(R_{l_M,a,t_p})$ je nach der Höhe des Alters a die entsprechende Relation aus Formel (3) einzusetzen hat.

---

1) Es wird vielleicht noch eine Rassenabhängigkeit des mittleren Milchertrags berücksichtigt werden müssen, was eine Spezifikation von Rassen-Unterlaufbahnen erfordern würde.

2) Der Altersbereich A ist hier durch l_M determiniert.

Das oben entwickelte Modell der Vorausrechnung von Erwartungswerten ist völlig identisch mit dem Modell, zu dem man gelangen würde, wenn man von der deterministischen Auffassung ausginge, dass die Schlachtungen nicht zufällig, sondern streng nach der Verteilung $s_1(\alpha)$ erfolgen, und dass auch die Zuordnung der Kälbergeburten in die Laufbahnen nicht zufällig, sondern streng nach der Verteilung γ_1 vorgenommen würden (was bedeutet, dass $s_1(\alpha)$ und γ_1 keine Wahrscheinlichkeiten, sondern feste Proportionen für diese Vorgänge angeben). Wegen dieser Gleichheit im Endeffekt können wir das obige Modell als deterministisches Modell ansprechen; gleichwohl unterscheidet es sich von einem rein deterministisch entstandenen darin, dass wir uns bei seiner Konzeption des stochastischen Charakters des Phänomens durchaus bewusst waren, denselben aber durch die Anwendung des Erwartungswertes als Sicherheitsäquivalent [1] ausgeschaltet haben.

Bevor wir zu einer Diskussion über den Wert eines solchen Modells schreiten, wollen wir uns zunächst noch mit den ihm unterstellten Voraussetzungen beschäftigen und uns überlegen, wie es um die praktische Bestimmbarkeit seiner Elemente bestellt ist.

3.2 Die im Modell implizierten Voraussetzungen

Das in 3.1 entwickelte Vorausrechnungsmodell enthält natürlich einige, die komplexe Realität des Phänomens vereinfachende Voraussetzungen, welche getroffen werden mussten, um das Problem mathematisch handhabbar zu machen. Wir wollen die wichtigsten dieser Voraussetzungen im folgenden diskutieren.

---

1) Vgl. hierzu [19], [28]. An späterer Stelle (vgl. §5) wird eine Rechtfertigung dieses Sicherheitsäquivalents gegeben werden.

Vorweg noch eine kurze Bemerkung. Die ganz zu Anfang getroffene Beschränkung der Betrachtung auf die Eigenentwicklung eines Prozesses ohne Berücksichtigung von Importen und Exporten stellt weniger eine Modellvoraussetzung als vielmehr eine Abgrenzung in Bezug auf den zu betrachtenden Gegenstand dar. Ebenso ist die Vernachlässigung der stochastischen Elemente nicht als eigentliche Voraussetzung, sondern als eine (vorläufige) Auffassung zu sehen. Wir werden diesen Punkt nachher ausführlicher erörtern.

Nun zu den eigentlichen Voraussetzungen des Modells.

Zunächst wurde beim Aufbau des Modells die zeitliche Stabilität[1] der Schlachtaltersverteilung $s_1(\alpha)$ und der Laufbahn-Verteilung γ_1 unterstellt. Auf lange Frist gesehen wäre dies eine fragwürdige Annahme. Da die Vorausrechnungen sich jedoch auf kurze Frist beschränken und zeitliche Stabilitäten als umso gesicherter angesehen werden können, je kürzer die Relevanzzeiträume sind, dürfte von dieser Seite her keine besonders schwerwiegende Fehlerquelle zu befürchten sein. Die eigentliche Rechtfertigung der zeitlichen Stabilitätsannahme besteht jedoch darin, dass sie eine absolute Notwendigkeit für das Modell darstellt, denn eine Voraussage der zeitlichen Entwicklung der Faktoren $s_1(\alpha)$ und γ_1 lässt sich ohne Berücksichtigung von prozessexogenen Faktoren nicht so leicht treffen und eine Extrapolation derselben aus der Vergangenheit heraus würde uns der Wirklichkeit auch nicht näher bringen. Kurz gesagt, die zeitliche Stabilitätsannahme kann zwar gewisse Tücken in sich bergen, sie ist aber nicht zu umgehen.

Es wurde ferner die regionale Stabilität der Modellkonstituenten vorausgesetzt, was allerdings bedenkenlos geschehen

1) Wegen näherer Betrachtungen über diese Problematik vgl. [21], [22].

konnte, da man im Falle der Nichthomogenität eines Gebietes hinsichtlich irgendwelcher Konstituenten das Gebiet jederzeit in homogene Regionen aufteilen und das Modell unter Berücksichtigung der jeweiligen Regionalwerte Region für Region anwenden kann. Als Modell für das gesamte Gebiet gilt dann die Summe der Regionalmodelle.

Bei einer praktischen Verwendung des Modells sollten sowohl die zeitliche als auch die regionale Stabilität von $s_1(\alpha)$ und γ_1 mittels geeigneter statistischer Prüfverfahren[1] laufend kontrolliert und an Hand solcher Kontrollen entsprechende Modifikationen des Modells bzw. seines Anwendungsmodus vorgenommen werden. Nur derartige gewissenhafte Prüfungen der Anwendungsvoraussetzungen können letztlich eine Gewähr für verlässliche Resultate bieten.

Eine weitere Voraussetzung des Modells besteht darin, dass zwischen den Laufbahnen 1 keine Uebergangsmöglichkeiten bestehen. Diese Annahme entspricht zwar nicht ganz der Wirklichkeit, aber zu ihrer Rechtfertigung lässt sich folgendes anführen : die Berücksichtigung von Uebergangsmöglichkeiten zwischen den Laufbahnen durch den Einbau von Markovschen Ketten[2] in das Modell würde dasselbe unverhältnismässig komplizieren und dabei keinen angemessenen Vorteil versprechen; der durch die Vernachlässigung dieses Phänomens begangene Fehler wird gering sein und kann als Störfaktor der Verteilung γ_1 zugeschlagen werden; zudem müssten bei Berücksichtigung von Uebergängen Markovsche Uebergangswahrscheinlichkeiten ermittelt werden, was auf nicht unerhebliche praktische Schwierigkeiten stossen würde. Die effiziente Vereinfachung des Modells und seiner Anwendung sprechen also für die Voraussetzung nicht vorhandener Uebergänge zwischen den Laufbahnen.

---

1) Bezüglich solcher Prüfverfahren vgl. z.B.[21],[22].

2) Zur Theorie der Markovschen Ketten vgl.[9], Seite 210 ff.

Die Voraussetzung des Fehlens von Laufbahnübergängen hat übrigens implizit die in §2.2 getroffene Vereinbarung zur Folge, nämlich dass alle Laufbahnen mit der Altersstufe 0 beginnen, dass also unmittelbar nach der Geburt eines Tieres die Entscheidung über seine Laufbahn fällt. Dies ist zwar individuell gesehen undenkbar, aber im Rahmen des Modells interessiert nur das globale Ergebnis einer Vielzahl solcher Entscheidungen, welches sich letztlich in der Laufbahnverteilung γ_1 manifestiert.

Die letzte noch zu erwähnende Voraussetzung des Modells ist die des Fehlens sonstiger Mortalität[1] (durch Unfall, Krankheit, Seuche u.s.w.). Diese Vereinfachung erfolgte im Hinblick darauf, dass die Notwendigkeit der Einbeziehung dieses Faktors in das Modell in jedem einzelnen Anwendungsfall erst untersucht werden muss. Sollte bei einem Prozess sonstige Mortalität in der Tat erhebliche Ausmasse annehmen, so ist noch ein Mortalitätsfaktor in das Modell einzubauen, und zwar als eine multiplikative Korrekturgrösse, welche abhängig ist vom Alter zum Zeitpunkt t_p und vom zeitlichen Abstand zwischen t_o und t_p.

Ein mehr oder weniger umfangreicher Voraussetzungskatalog wie der obige liegt explizit oder implizit jedem Modell eines Phänomens unserer Umwelt zugrunde. Derartige Voraussetzungen idealisieren immer die Wirklichkeit und führen zu einem verfälschten Bild derselben. Die so entstehende Abweichung des Modells von der Realität bezeichnet man als den

---

1) Darunter ist solche Mortalität zu verstehen, welche nicht auf Schlachtung beruht.

Spezifikationsfehler des Modells. Für das Ausmass dieses Spezifikationsfehlers, welcher wiederum massgeblich die Zuverlässigkeit der mit dem Modell erzielten Ergebnisse beeinflusst, ist also die Realitätsangemessenheit der Voraussetzungen bestimmend, und es liegt somit im Interesse des Anwendungserfolges eines Modells, die Voraussetzungen so gewissenhaft wie möglich zu prüfen. In vielen Fällen bieten sich hierzu geeignete statistische Prüfverfahren an. Bei unzulänglichen Resultaten eines Modells ist neben der Zuverlässigkeit der Modelldaten gleichzeitig immer auch der Voraussetzungskatalog zu überprüfen.

3.3 Verfügbarkeit bzw. Bestimmbarkeit der Modelldaten

Grundvoraussetzung für die praktische Anwendung des in 3.1 entwickelten Modells ist selbstverständlich die Verfügbarkeit bzw. Bestimmbarkeit der Daten R_{t_o}, $s_1(\alpha)$, γ_1 und K_t. Wir wollen dieses Problem kurz erörtern.[1)]

Der Bestand R_{t_o} : Der Bestand zum Basiszeitpunkt t_o kann unmittelbar zur Verfügung stehen, wenn gerade in t_o eine Bestandserhebung vorliegt. Wenn die letzte Bestandserhebung soweit zurückliegt, dass sie nicht mehr als aktuell gelten kann, so kann, wenn entsprechende Angaben über Neuzugänge und Schlachtungen zwischen jenem letzten Erhebungszeitpunkt und t_o vorliegen, mittels der Fortschreibungsformel (1) der Bestand auf den in t_o aktuellen Stand gebracht werden.

---

1) Auf eine Behandlung der sich stellenden schätztheoretischen Probleme verzichten wir hier, werden aber an späterer Stelle (§ 4.2) darauf zu sprechen kommen. Die Aspekte der Stichprobentheorie und Stichprobentechnik werden hier ausser Acht gelassen.

Die Schlachtaltersverteilung $s_1(\alpha)$:

Die Bestimmung der Schlachtaltersverteilung kann an Hand von Kohortenanalysen[1] erfolgen, welche sich aus Buchführungsdaten und Prozessbeobachtungen rekonstruieren lassen. Mit den Kohortenanalysen sind Bestandsabnahmekurven gleichaltriger Tiere zu ermitteln, aus deren durchschnittlichem Verhalten mit Hilfe der Relation (2) entsprechende Schlachtaltersverteilungen hergeleitet werden können. Die Resultate liefern erste empirische Approximationen der "wahren" Schlachtaltersverteilung. Dieselben bedürfen vermutlich noch einer Bereinigung von zufälligen Störungseinflüssen, ehe sie eine analytische Darstellung erhalten können.

Die Laufbahnverteilung γ_1 :

Für die Bestimmung der Verteilung γ_1 ist ebenfalls die Kohortenanalyse das geeignete Instrument. Es ist dabei aus Prozessbeobachtungen bzw. Buchführungsdaten bei Viehhaltern und Schlachthäusern festzustellen, welcher Anteil von geborenen Kälbern in welcher Laufbahn endet.

Die Kälbergeburten K_t :

Die Bestimmung der Kälbergeburten K_t für die einzelnen Punkte t des Zeitraumes (t_o, t_p) stellt ein eigenes Vorausrechnungsproblem im Rahmen des Vorausrechnungsmodells dar. Wir wollen 2 Wege angeben, welche bei diesem Problem zum Ziele führen können :

a) Zeitreihenanalyse :

Eine erste Möglichkeit, eine Vorausrechnung der Kälbergeburten K_t für die einzelnen Punkte des Zeitraums (t_o, t_p) vorzunehmen, liefern die Methoden der

1) Vgl. hierzu [34], Seite 34-37.

Zeitreihenanalyse[1]. Man hat dabei im Prinzip so vorzugehen, dass man aus der Vergangenheit bekannte lange Zeitreihen über Kälbergeburten unter Eliminierung von Zufallseinflüssen in die Komponenten Trend und Saisoneinfluss zerlegt und dieselben bestimmt. Indem man Trend und Saisoneinfluss getrennt extrapoliert und dieselben für jeden gewünschten Zeitpunkt $t \in (t_o, t_p)$ wieder zusammenfügt, erhält man für den betreffenden Zeitpunkt eine Vorausrechnung für die Kälbergeburten, welche in das Vorausrechnungsmodell einzufügen ist.

b) Regressionsanalyse :

Eine zweite Möglichkeit, die Kälbergeburten vorauszurechnen, besteht in einer Regressionsanalyse zwischen den Kälbergeburten K_t und dem Milchkuhbestand $R_{1_M, t-\tau}$[2] zum früheren Zeitpunkt $t - \tau$.

Zunächst bestimmt man dazu die Zeitverzögerung τ an Hand von früheren Beobachtungen so, dass die Korrelation zwischen K_t und $R_{1_M, t-\tau}$ maximal wird. Erreicht der Korrelationskoeffizient einen hohen Wert, so besteht zwischen K_t und $R_{1_M, t-\tau}$ ein enger linearer Zusammenhang der Form

$$K_t = a_o + a_1 R_{1_M, t-\tau} + \varepsilon_t \quad (8)$$

[3]

wobei ε_t eine stochastische Störvariable darstellt.

Dieser Zusammenhang soll dazu benutzt werden, K_t vorauszurechnen.

Dazu werden zunächst die Koeffizienten a_o und a_1 an Hand früherer Beobachtungen zweckmässigerweise mittels der

---

1) Vgl. hierzu [4], [10].

2) $R_{1M, t-\tau}$ ist gleich $\int_A R_{1M, a, t-\tau} da$ (vgl. hierzu Formel (7)).

3) Eventuell ist ein multipler Regressionsansatz mit mehreren Verzögerungen $\tau_1, \ldots, \tau_k$ wirklichkeitsnäher. Zur Regressionsanalyse im allgemeinen vgl. man z.B. [10], [16].

Kleinst-Quadrat-Methode[1] geschätzt. Mit vorliegenden Schätzungen $\hat{a}_o$ und $\hat{a}_1$ für a_o und a_1 kann dann vermöge der Relation (8) die Grösse K_t durch die um τ verzögerte (vorherbestimmte) Grösse $R_{l_M, t-\tau}$ substituiert werden. Der Ausdruck $\hat{a}_o + \hat{a}_1 R_{l_M, t-\tau}$ liefert dann die gewünschte Vorausrechnung für K_t.

Ein besonderes Problem bei diesem Vorgehen stellt sich dann, wenn $t - \tau$ einen zukünftigen Zeitpunkt darstellt und $R_{l_M, t-\tau}$ also keine bekannte Grösse mehr ist. In diesem Fall hat man zunächst $R_{l_M, t-\tau}$ mittels Relation (3) aus dem letzten bekannten Bestand vorauszuberechnen, um schliesslich diesen vorausgerechneten Wert für $R_{l_M, t-\tau}$ in den Ausdruck $\hat{a}_o + \hat{a}_1 R_{l_M, t-\tau}$ einzusetzen.

Eine fundamentale praktische Schwierigkeit bei einem solchen Regressionsverfahren wird sich jedoch daraus ergeben, dass Bestandserhebungen wegen ihrer hohen Kosten vermutlich nur in grösseren Zeitabständen erfolgen, so dass man nicht über die erforderlichen laufenden Angaben über $R_{l_M, t}$ verfügen würde.

Angesichts dieser Schwierigkeit könnte man die folgende Möglichkeit einer Vorausrechnung für K_t ins Auge fassen. Anstatt mit der Regression zwischen K_t und $R_{l_M, t-\tau}$ könnte man auf der Basis der Anzahl D_{t-9} der 9 Monate vor t gedeckten oder befruchteten Tiere zu operieren versuchen. Hier bestünde zwar bei einer Informationsbeschaffung über D_t in Verbindung mit Bestandserhebungen ebenfalls das Problem zeitlich zu lückenhafter

---

1) Wegen Einzelheiten darüber vgl. [10], [16].

Angaben, aber im Prinzip liessen sich laufende und lückenlose Angaben durch laufende, von den Bestandserhebungen unabhängige Deckungs- und Befruchtungsstatistiken[1] gewinnen. Vorausrechnungen für K_t wären auf diese Art natürlich nur bis maximal 9 Monate nach dem Gegenwartszeitpunkt möglich, darüber hinaus hätte man wieder mit vorausberechneten Beständen $R_{1_M,t}$ zu arbeiten.

Es wäre hier zu erwähnen, dass z.B. auf dem Schweinesektor in Frankreich und den Niederlanden Deckungsstatistiken geführt werden, aus welchen sich mit Hilfe von Schätzungen über durchschnittliche Wurfgrössen die Anzahl der Ferkelgeburten zu entsprechend späterem Zeitpunkt ziemlich genau vorausberechnen lassen.

3.4 Wert und Grenzen dieses deterministischen Modells

Das obige deterministische Modell liefert als Information über das künftige Ereignis einen einzigen Vorausrechnungswert ohne irgendeine Angabe über die Zuverlässigkeit desselben. Wann ist ein solches Modell von Nutzen?

---

1) In Verbindung mit laufenden Schlachtungsstatistiken könnte übrigens eine solche laufende Deckungs- und Befruchtungsstatistik die kostspieligeren Bestandserhebungen bis zu einem gewissen Grad ersetzen. Man könnte nämlich den zu erhebenden Bestand an Hand eines früher erhobenen Bestandes und den laufenden Angaben über Deckungen (und Befruchtungen) sowie über Schlachtungen mittels einer leicht modifizierten Form der Fortschreibungsformel (1) ermitteln. Diese Modifikation von (1) würde nur darin bestehen, dass man die Grösse N_{1t} durch die Grösse $\gamma_1 \cdot f(D_{t-9})$ ersetzt, wobei die Funktion f die Regressionsfunktion zwischen K_t und D_{t-9} ist.

Solche Modelle sind in zwei Fällen anwendbar:

1. Wenn das betrachtete Phänomen selbst deterministischer Natur ist und auf Grund exakter mathematischer Beziehungen zwischen erklärenden und zu erklärenden Grössen gar keine oder nur ganz unbedeutende Abweichungen zwischen Vorausrechnung und Wirklichkeit auftreten können.

   Solchen Phänomenen begegnet man eigentlich nur in der toten Natur. Bei unserem Problem des Rinderprozesses manifestiert sich die nicht deterministische Natur bereits eindeutig in der Schlachtaltersverteilung $s_1(\alpha)$ und der Laufbahnverteilung γ_1. Vom Phänomen her gesehen besteht also keine Rechtfertigung der Anwendung eines deterministischen Modells.[1)]

2. Wenn Vorausrechnungen für anspruchslose informationsextensive Belange gewünscht werden.

   Solche Belange können auf dem Rindersektor beispielsweise darin bestehen, dass man auf Grund einer Vorausrechnung lediglich die Tendenz einer Entwicklung frühzeitig zu erkennen wünscht und im Hinblick auf dieselbe vielleicht grobe regionale Vergleiche anzustellen gedenkt. Es können also durchaus Verwendungszwecke für Vorausrechnungen vorliegen, bei welchen Genauigkeitsangaben keine besondere Bedeutung haben, zumal wenn man vielleicht weiss, dass die angewendete Vorausrechnungsmethode eine bestimmte Mindestgenauigkeit garantiert. Bei solchen Nachfragen kann das obige Modell nützliche Dienste leisten.

Ein informationsintensiver Bedarf für Vorausrechnungen dagegen, wie er sich etwa ergibt im Zusammenhang mit der Vorbereitung von Entscheidungen von grosser Tragweite, kann

1) Wenn bei dem Phänomen allerdings ausser den von $s_1(\alpha)$ und γ_1 beschriebenen Ungewissheiten keine weiteren Ungewissheiten auftreten, so liefert, wenn die Bestände und Kälbergeburten sehr hoch sind, das Gesetz der grossen Zahl doch eine Rechtfertigung der deterministischen Auffassung. Wir werden darauf am Ende des §6 zu sprechen kommen.

von diesem deterministischen Modell nicht befriedigt werden. Um beispielsweise im Rahmen einer rationellen Marktpolitik optimale Entscheidungen zu treffen, sind Genauigkeitsangaben für Vorausrechnungen unentbehrlich, da sie es erst ermöglichen, das mit den Entscheidungsalternativen verbundene Risiko zu messen, an welchem man letztlich die Entscheidung orientiert. Angesichts der materiellen oder ideellen Konsequenzen bei Fehlentscheidungen wird man vielleicht noch abzuwägen versuchen, wie weit es sich lohnt, Mittel für genauere Vorausrechnungen bereitzustellen. Solche Ueberlegungen können offenbar nicht auf der Grundlage des obigen Modells angestellt werden. Die sich bei derartigen Untersuchungen stellenden Ansprüche an Vorausrechnungen überfordern das obige Modell eindeutig.

Um auf solche Probleme eingehen zu können, was wir im Rahmen dieser Arbeit noch zu tun gedenken, bedarf es also der Entwicklung eines Modells, welches über eine solchen Vorhaben angemessene Ausstattung verfügt.

Zu einem solchen Modell gelangen wir, wenn wir in die Konstruktionsidee des obigen deterministischen Modells den stochastischen Charakter der einzelnen Konstituenten mit einbeziehen. Wir werden uns jetzt dieser Aufgabe zuwenden.

§ 4 EIN STOCHASTISCHES MODELL ZUR VORAUSRECHNUNG EINES RINDERPROZESSES

4.1 Das Modell

Das oben dargestellte deterministische Modell werden wir jetzt durch ein stochastisches Modell ablösen, welches wir dann unseren weiteren Betrachtungen zugrunde legen. Wir beschränken uns dabei auf die Darstellung des der Formel(3) entsprechenden stochastischen Modells, also des Modells zur Vorausrechnung von R_{lat_p}.

Das stochastische[1)] Modell soll den Zusammenhang herstellen zwischen der zu erklärenden unbekannten Grösse R_{lat_p} und den als stochastische Variablen interpretierten erklärenden

---

1) Von einem stochastischen Modell sprechen wir aus folgendem Grunde : Die Grundauffassung, auf welcher wir das Modell aufbauen ist die, dass alle mit Ungewissheit behafteten Grössen als sich im Zustand stochastischer Veränderlicher befindend anzusehen sind; es wird also davon ausgegangen, dass alle auftretenden Ungewissheiten mit Wahrscheinlichkeitsverteilungen gemessen werden. Diese Auffassung entspricht übrigens der der Bayesschen Schule (vgl.[11],[26]), nach welcher unbekannte Grössen immer als stochastische Veränderliche zu betrachten sind.

Diese Auffassung ist für unser Problem aus folgendem Grund zweckmässig : In dem stochastischen Modell soll, ganz allgemein gesprochen, die Ungewissheit einer zu erklärenden Grösse Y, welche durch eine bekannte Funktion von anderen ungewissen Grössen X_i abhängt, durch die Ungewissheiten dieser Grössen ausgedrückt werden. Dies ist aber prinzipiell wegen gewisser mathematischer Erfordernisse bezüglich des verwendeten Ungewissheitsmaßes (dasselbe muss die Eigenschaft der Additivität erfüllen) nur möglich, wenn die Ungewissheiten in Wahrscheinlichkeiten ausgedrückt werden. Bei anderen Ungewissheitsmaßen wie den Konfidenzaussagen oder der Likelihoodfunktion z.B. ist es nicht möglich von Aussagen über zwei Grössen X_1 und X_2 Aussagen über eine Funktion der beiden,wie z.B. $X_1 + X_2$,herzuleiten (vgl.[24],Anhang A.3.3.4).

Grössen $R_{1,a-t_p+t_o,t_o}$, K_{t_p-a}, $s_1(\alpha)$ und γ_1 [1].
Es soll ferner zeigen, wie sich die Wahrscheinlichkeitsverteilung von R_{lat_p} aus den Wahrscheinlichkeitsverteilungen dieser erklärenden Grössen zusammensetzt.

Zunächst gehen wir davon aus, dass alle Grössen $R_{1,a-t_p+t_o,t_o}$, K_{t_p-a}, $s_1(\alpha)$ und γ_1 exakt bekannt sind. Der stochastische Charakter von R_{lat_p} wird dann nur von den durch $s_1(\alpha)$ und γ_1 beschriebenen Ungewissheiten geprägt.

Das Zustandekommen von R_{lat_p} lässt sich dann folgendermassen als das Ergebnis eines stochastischen Experimentes deuten :

Für $a \geq t_p - t_o$ ist R_{lat_p} gleich der Summe der Ergebnisse unabhängiger (0,1) - Experimente, deren je eines man sich mit jedem der $R_{1,a-t_p+t_o,t_o}$ Tiere durchgeführt vorstellt. Dabei bedeutet das Ereignis 1, dass das Tier zum Zeitpunkt t_p noch lebt, wofür die Wahrscheinlichkeit gegeben ist durch

$$p_{lat_p t_o} \overset{def}{=} \frac{\int_a^\infty s_1(\alpha)d\alpha}{\int_{a-t_p+t_o}^\infty s_1(\alpha)d\alpha} .$$

Das Ereignis 0 bedeutet, dass das Tier vor dem Zeitpunkt t_p geschlachtet wird, wofür die Wahrscheinlichkeit gleich $1 - p_{lat_p t_o}$ ist.

---

1) Unter stochastisch variablen Verteilungen $s_1(\alpha)$ und γ_1 ist zu verstehen, dass $s_1(\alpha)$ und γ_1 für jedes feste α bzw. 1 stochastische Grössen sind.

Für $a < t_p - t_o$ hat man sich diese Experimentserie zweimal hintereinander durchgeführt vorzustellen.

In der ersten Serie wird mit jedem der K_{t_p-a} Tiere, welche zum Zeitpunkt $(t_p - a)$ zur Welt kommen, je ein von den anderen unabhängiges (0,1)-Experiment durchgeführt. Dabei bedeutet das Ereignis 1, dass das Tier in die Laufbahn 1 kommt, wofür die Wahrscheinlichkeit gleich γ_1 ist. Das Ereignis 0 bedeutet, dass das Tier nicht in die Laufbahn 1 kommt, wofür die Wahrscheinlichkeit gleich $1 - \gamma_1$ ist.

Die (stochastische) Summe der Ergebnisse dieser Experimentreihe gibt an, wieviele neugeborene Tiere sich zum Zeitpunkt $(t_p - a)$ in der Laufbahn 1 befinden (nämlich R_{1,o,t_p-a}).

R_{1at_p} ergibt sich schliesslich daraus, dass man mit jedem der $R_{1,o,t_p - a}$ Tiere sich wieder ein unabhängiges (0,1)-Experiment durchgeführt vorstellt, bei welchem das Ereignis 1 bedeutet, dass das Tier in t_p noch lebt (wofür die Wahrscheinlichkeit gleich

$$q_{1a} \overset{def}{=} \int_a^\infty s_1(\alpha)d\alpha \quad \text{ist)}$$

und wobei 0 das Ereignis ist, dass das Tier im Zeitraum $(t_p - a, t_p)$ geschlachtet wird (wofür die Wahrscheinlichkeit gleich $1 - q_{1a}$ ist).

Mathematisch lässt sich dies folgendermassen ausdrücken :

(9) $R_{lat_p} =$

$$\begin{cases} \displaystyle\sum_{i=1}^{R_{1,a-t_p+t_o,t_o}} z_i & \text{für} \quad a \geq t_p - t_o, \\ \qquad \text{wobei die } z_i \text{ voneinander unabhängige (0,1)-Variablen sind mit 1-Wahrscheinlichkeit} \\ \qquad p_{lat_p t_o} = \int_a^\infty s_1(\alpha)d\alpha \Big/ \int_{a-t_p+t_o}^\infty s_1(\alpha)d\alpha \,. \\ \displaystyle\sum_{i=1}^{R_{1,o,t_p-a}} z_i^* & \text{für} \quad a < t_p - t_o, \\ \qquad \text{wobei die } z_i^* \text{ voneinander unabhängige (0,1)-Variablen sind mit 1-Wahrscheinlichkeit} \\ \qquad q_{1a} = \int_a^\infty s_1(\alpha)d\alpha \,; \\ \qquad \text{die Grösse } R_{1,o,t_p-a} \text{ ist ebenfalls eine stochastische Veränderliche, welche sich darstellen lässt durch} \\ \qquad R_{1,o,t_p-a} = \sum_{j=1}^{K_{t_p-a}} z_j^{**} \,, \\ \qquad \text{wobei die } z_j^{**} \text{ voneinander unabhängige (0,1)-Variablen sind mit 1-Wahrscheinlichkeit } \gamma_1. \end{cases}$$

Die Wahrscheinlichkeitsverteilung von R_{lat_p}, welche wir mit $\omega(R_{lat_p})$ bezeichnen wollen, ist demnach für $a \geq t_p - t_o$ eine Binomialverteilung. Für $a < t_p - t_o$ dagegen handelt es sich um eine Verteilung, welche sich aus einer Art Ueber-

lagerung zweier Binomialverteilungen ergibt. Die Verteilung $\omega(R_{lat_p})$ ist gegeben durch:

$$
(10)\quad \begin{cases} pr(R_{lat_p}=k \mid R_{1,a-t_p+t_o,t_o}, p_{lat_p t_o}) \\ = \binom{R_{1,a-t_p+t_o,t_o}}{k} p_{lat_p t_o}^{k} (1-p_{lat_p t_o})^{R_{1,a-t_p+t_o,t_o}-k}, \quad a \geq t_p - t_o \\ pr(R_{lat_p}=k \mid K_{t_p-a}, q_{1a}, \gamma_1) \\ = \sum_{n=o}^{K_{t_p-a}} (\tfrac{n}{k}) \binom{K_{t_p-a}}{n} q_{1a}^{k} (1-q_{1a})^{n-k} \gamma_1^{n} (1-\gamma_1)^{K_{t_p-a}-n}, \quad a < t_p - t_o. \end{cases}
$$

Bei der Ermittlung der Verteilung für $a<t_p-t_o$ hat man die Beziehung $f(x)=\int_Y f_1(x \mid y) f_2(y) dy$ zwischen marginalen und bedingten Verteilungen zu beachten.

Wir waren oben zunächst von der Annahme ausgegangen, dass alle Bestimmumgsgrössen von R_{lat_p} exakt gegebene Grössen sind. Dies ist selbstverständlich eine unrealistische Annahme, denn in der Praxis kommen alle diese Grössen durch Schätzungen bzw. Vorausrechnungen zustande, so dass sie alle, einschliesslich der Verteilungen $s_1(\alpha)$ und γ_1, mit Ungewissheiten behaftet sind. Um diesem Umstand Rechnung zu tragen, werden wir das Modell entsprechend erweitern.

Wir gehen jetzt davon aus, dass die erklärenden Grössen von R_{lat_p} selbst ungewiss und damit im weiteren (Bayesschen) Sinn stochastische Grössen sind. Wir nehmen aber auch an, dass wir über Informationen verfügen können, welche es uns ermöglichen, Wahrscheinlichkeitsverteilungen für diese Grössen zu konstruieren. Wie dies geschehen kann, werden

wir nachher zeigen. Wir gehen jetzt davon aus, dass, je nach Informationslage, gewisse Wahrscheinlichkeitsverteilungen φ_i für die erklärenden Grössen gegeben sind und erweitern das Modell (9) um den folgenden Zusatz :

(11)
- Für $a \geq t_p - t_o$ nehmen wir an, dass $R_{1,a-t_p+t_o,t_o}$ und $p_{lat_p t_o}$ unabhängige stochastische Veränderliche mit bestimmten Wahrscheinlichkeitsverteilungen $\varphi_1(R_{1,a-t_p+t_o,t_o})$[1] und $\varphi_2(p_{lat_p t_o})$ sind.
- Für $a < t_p - t_o$ setzen wir voraus, dass K_{t_p-a}, q_{1a} und γ_1 unabhängige stochastische Veränderliche mit bestimmten Wahrscheinlichkeitsverteilungen $\varphi_3(K_{t_p-a})$, $\varphi_4(q_{1a})$ und $\varphi_5(\gamma_1)$ sind.

Das um diesen Zusatz (11) ergänzte Modell (9) stellt das stochastische Modell dar, auf welchem wir unsere weiteren Untersuchungen aufbauen wollen.

Die Wahrscheinlichkeitsverteilungen in (10) stellen in Bezug auf das Modell (9,11) nur bedingte Verteilungen für R_{lat_p} dar. Um aus diesen bedingten Verteilungen (10) die Marginalverteilungen von R_{lat_p} zu bilden, hat man dieselben mit den Verteilungen φ_i der jeweiligen Bedingungen zu multiplizieren und anschliessend über diese Bedingungen zu integrieren. Die Verteilung $\omega(R_{lat_p})$ ist dann gegeben durch :

---

1) Die Verteilungen φ_i hängen natürlich von denselben Indizes ab wie ihre Argumente (d.h. φ_1 z.B. hat man sich noch mit den Indizes $1,a-t_p+t_o,t_o$ versehen vorzustellen).

$$(12)\quad \begin{cases} \mathrm{pr}(R_{lat_p}=k) = \int_p \int_R \binom{R}{k} p^k (1-p)^{R-k} \varphi_1(R)\varphi_2(p)\,dR\,dp & \text{für } a \geq t_p - t_o \\ \mathrm{pr}(R_{lat_p}=k) & \text{für } a < t_p - t_o \\ = \int_\gamma \int_q \int_K \sum_{n=o}^{K} \binom{n}{k}\binom{K}{n} q^k (1-q)^{n-k}\, \gamma^n (1-\gamma)^{K-n}\, \varphi_3(K)\varphi_4(q)\varphi_5(\gamma)\,dK\,dq\,d\gamma\,. \end{cases}$$

Die der Uebersichtlichkeit wegen vernachlässigten Indizes von R, p, K, q und γ sind dieselben wie in (10).

Die Ausdrücke (12) würden also die Wahrscheinlichkeitsverteilung $\omega(R_{lat_p})$ angeben. Es ist offensichtlich, dass selbst unter sehr vereinfachenden Annahmen bezüglich der Verteilungen φ_i die Verteilung von R_{lat_p} keine sehr gefällige analytische Form annehmen wird. Wir verzichten auf weitere Untersuchungen in dieser Hinsicht und begnügen uns damit zu wissen, dass wir mit Hilfe von (12) im Prinzip die Verteilung von R_{lat_p} bei gegebenen Verteilungen φ_i bestimmen könnten.

Nachher wird sich sogar herausstellen, dass man gar nicht immer auf die Kenntnis der gesamten Verteilung $\omega(R_{lat_p})$ angewiesen ist, sondern dass unter bestimmten Gesichtspunkten gewisse Parameter von $\omega(R_{lat_p})$ alle relevanten Informationen über R_{lat_p} enthalten. Man wird sich also auf diese Parameter beschränken und dieselben dann aus den entsprechenden Parametern der Verteilungen φ_i zu ermitteln suchen.

Abschliessend wollen wir noch darauf hinweisen, dass sämtliche in § 3.2 aufgezählten Voraussetzungen hier ebenfalls unterstellt wurden. Als zusätzliche plausible

Voraussetzung hatten wir in (11) noch die stochastische Unabhängigkeit gewisser Grössen gefordert.

4.2 Die Konstruktion von Wahrscheinlichkeitsverteilungen φ_i für die erklärenden Grössen des Modells

Wir werden jetzt untersuchen, wie man aus gewissen Informationen Wahrscheinlichkeitsverteilungen φ_i für die Grössen $R_{l,a-t_p+t_o,t_o}$, $p_{lat_p\bar{t}_o}$, K_{t_p-a}, q_{la} und γ_l konstruieren kann. Dies ist allgemein gesprochen das Problem der statistischen Inferenz[1].

Zunächst wollen wir eine geeignete Inferenzmethode wählen. Wir haben dabei die folgende Auswahl[2]:

a) die Bayessche Methode
b) die Fiduzial-Methode
c) die Konfidenz-Methode
d) die Likelihood-Methode.

Die Konfidenz-Methode und die Likelihood-Methode scheiden sofort aus, da sie der oben getroffenen Vereinbarung, Ungewissheiten nur mit Wahrscheinlichkeitsverteilungen zu messen, nicht genügen können[3]. Auch auf die Fiduzial-Methode wollen wir wegen ihrer sehr begrenzten Anwendbarkeit verzichten[4], so dass uns schliesslich nur noch die Bayessche Methode verbleibt. Die Bayessche Methode besteht in folgender Schlussweise :

$f(x;\theta)$ sei die Wahrscheinlichkeitsverteilung einer stochastischen Veränderlichen x in Abhängigkeit vom unbekannten

---

1) Vgl. hierzu [18], Seite 265 ff.
2) Wegen näherer Einzelheiten über diese Methoden vgl.[18],[29].
3) Vgl. hierzu Fussnote 1 Seite 35.
4) Vgl.[18], Seite 279.

Parameter Θ. Wenn über x eine Beobachtung $x=x_o$ vorliegt und für Θ eine a-priori-Verteilung $p(\Theta)$ gegeben ist, dann ist durch das Bayessche Theorem in Form der a-posteriori-Verteilung

$$(13) \qquad p_B(\Theta|x_o) = \frac{f(x_o;\Theta)p(\Theta)}{\int_\Theta f(x_o;\Theta)p(\Theta)d\Theta}$$

eine durch die Information x_o induzierte Wahrscheinlichkeitsverteilung für Θ gegeben. Diese Verteilung setzt sich also zusammen aus der a-priori-Verteilung $p(\Theta)$ und der sogenannten Likelihoodfunktion $L(\Theta|x_o)=f(x_o;\Theta)$ für Θ.

Wir werden jetzt überlegen, wie wir dieses Prinzip zur Konstruktion von Verteilungen φ_i aus gegebenen Informationen anwenden können.

Wir beginnen mit der Konstruktion einer Verteilung $\varphi_5(\gamma_1)$:

Im Rahmen einer Kohortenanalyse habe man eine Beobachtung gemacht, dass von n neugeborenen Kälbern k in die Laufbahn 1 gelangen. Wie können wir aus dieser Information eine Wahrscheinlichkeitsverteilung für die Grösse γ_1 konstruieren?

Wenn γ_1 die Wahrscheinlichkeit dafür ist, dass ein Tier in die Laufbahn 1 eingeordnet wird, dann ist die Wahrscheinlichkeit der Beobachtung, dass von n Tieren k in die Laufbahn 1 kommen, gleich $\binom{n}{k}\gamma_1^k(1-\gamma_1)^{n-k}$.

Wenn wir noch annehmen, dass die Grösse γ_1 im Intervall $[0,1]$ als a-priori-Verteilung eine Gleichverteilung [1] besitzt, dann ist nach (13) die a-posteriori-Verteilung für γ_1 gegeben durch :

---

1) Diese Annahme ist übrigens als das Bayessche Postulat bekannt (vgl.[9], Seite 416).

$$(14) \qquad p_B(\gamma_1 | n,k) = \frac{\gamma_1^k(1-\gamma_1)^{n-k}}{\int_0^1 \gamma_1^k(1-\gamma_1)^{n-k}d\gamma_1} \quad , \qquad \gamma_1 \in [0,1].$$

Der Ausdruck (14) liefert uns also eine auf der Beobachtung (n,k) beruhende Verteilung $\varphi_5(\gamma_1)$.

Genau in der gleichen Weise wie für γ_1 kann man auch verfahren, um aus gewissen Beobachtungen Wahrscheinlichkeitsverteilungen für q_{1a} und $p_{1at_pt_o}$ zu gewinnen.

Zur Konstruktion einer Verteilung $\varphi_1(R_{1,a-t_p+t_o,t_o})$:

Zur Herleitung einer Wahrscheinlichkeitsverteilung für den Bestand $R_{1,a-t_p+t_o,t_o}$ zerlegen wir denselben (der Einfachheit wegen unter Vernachlässigung der Indizes)in

$$R = N \cdot r$$

wobei wir N als die Anzahl von Erhebungseinheiten (z.B. die landwirtschaftlichen Betriebe mit Rinderhaltung) und r als den mittleren Bestand pro Erhebungseinheit betrachten.

Mit x bezeichnen wir schliesslich die Anzahl der Tiere pro Betrieb. Die Grösse x besitze eine Verteilung mit Mittelwert r und bekannter Streuung σ^2. Das arithmetische Mittel $\bar{x}$ aus einer genügend grossen Anzahl unabhängiger Beobachtungen $x_1, \ldots, x_k$ besitzt dann nach dem zentralen Grenzwerttheorem[1] eine Normalverteilung mit Mittelwert r und Streuung $\frac{\sigma^2}{k}$. Wenn dann eine aus k Einzelbeobachtungen gebildete Mittelwertbeobachtung $\bar{x}_o$ vorliegt, und wir für r als a-priori-Verteilung eine Gleichverteilung in dem für r vorgegebenen Intervall $[o,a]$, $a > o$, annehmen, dann ist gemäss (13) die a-posteriori-Verteilung für r gegeben

1) Vgl. hierzu [9], Seite 174.

durch

$$(15)\quad p_B(r|\bar{x}_o)=\begin{cases}\dfrac{\frac{1}{\sigma}(k/2\pi)^{\frac{1}{2}}\exp[-\frac{k(r-\bar{x}_o)^2}{2\sigma^2}]\frac{1}{a}}{\int_0^a \frac{1}{\sigma}(k/2\pi)^{\frac{1}{2}}\exp[-\frac{k(r-\bar{x}_o)^2}{2\sigma^2}]\frac{1}{a}\,dr} & \text{für } r\in[o,a]\\ 0 & \text{für } r \text{ ausserhalb}[o,a].\end{cases}$$

Wenn wir a als unendlich gross auffassen (d.h. r ist über der ganzen nichtnegativen Achse gleichverteilt) und für (15) den Grenzübergang "a gegen Unendlich" bilden, so erhalten wir

$$(16)\quad p_B^*(r|\bar{x}_o)=\begin{cases}\dfrac{\exp[-k(r-\bar{x}_o)^2/2\sigma^2]}{\int_0^\infty \exp[-k(r-\bar{x}_o)^2/2\sigma^2]dr} & \text{für } r\geq o\\ 0 & \text{für } r< o.\end{cases}$$

Diese Wahrscheinlichkeitsverteilung stellt eine am Nullpunkt gestutzte[1] Normalverteilung dar.

Aus (16) erhalten wir schliesslich für $R = Nr$ die Verteilung

$$(17)\quad p_B^{**}(R|\bar{x}_o)=\begin{cases}\dfrac{\exp[-k(R-N\bar{x}_o)^2/2N^2\cdot\sigma^2]}{\int_0^\infty \exp[-k(R-N\bar{x}_o)^2/2N^2\cdot\sigma^2]dR} & \text{für } R\geq o\\ 0 & \text{für } R< o.\end{cases}$$

Die Verteilung (17) stellt also eine Verteilung φ_1 dar. Mittelwert und Streuung dieser Verteilung stimmen für den Fall, dass die links des Nullpunkts von der Normalverteilung abgeschnittene Wahrscheinlichkeitsmasse sehr klein ist, annähernd mit $N\bar{x}_o$ bzw. $N^2\cdot\sigma^2/k$ überein.

---

1) Zum Begriff der gestutzten (engl. truncated) Verteilungen vgl.z.B.[18], Seite 201 f.

Zur Konstruktion einer Verteilung $\varphi_{\beta}(K_{t_p-a})$:

Wir beziehen uns hierzu auf den Regressionsansatz (3) in § 3.3. Die Konstruktion einer Wahrscheinlichkeitsverteilung für K_t in Abhängigkeit von einem gegebenen $R_{l_M,\, t-\tau}$ aus früheren Beobachtungen $(K_t, R_{l_M,t-\tau})$, $t = 1, 2, \ldots, n$, stellt uns im Prinzip vor das folgende Problem :

Der stochastische Zusammenhang zwischen einer Grösse y und r Grössen $x_1, \ldots, x_r$, welcher auch noch von k unbekannten Parametern $\theta_1, \ldots, \theta_k$ mitbestimmt wird, sei beschrieben durch eine Verteilung $f(y;\theta_1, \ldots, \theta_k, x_1, \ldots, x_r)$.

Es liegen n unabhängige Beobachtungen $z_i = (y_i; x_{1i}, \ldots, x_{ri})$ $(i=1,\ldots,n)$ vor und es sei ferner eine a-priori-Verteilung $p(\theta_1, \ldots, \theta_k)$ für die unbekannten Parameter gegeben. Wie kann man daraus eine zu vorgegebenen Grössen $(x_{1o}, \ldots, x_{ro})$ gehörige Wahrscheinlichkeitsverteilung für y angeben?

Es sei

$$(18) \quad L(\theta_1, \ldots, \theta_k | z_1, \ldots, z_n) = \prod_{i=1}^{n} f(y_i; \theta_1, \ldots, \theta_k, x_{1i}, \ldots, x_{ri})$$

die Likelihood-Funktion von $(\theta_1, \ldots, \theta_k)$ bei gegebenen Beobachtungen $(y_i; x_{1i}, \ldots, x_{ri})$, $i=1, \ldots, n$.

Dann erhält man gemäss (13) eine a-posteriori-Verteilung

$$(19) \quad p_B(\theta_1, \ldots, \theta_k | z_1, \ldots, z_n) = \frac{L(\theta_1, \ldots, \theta_k | z_1, \ldots, z_n) p(\theta_1, \ldots, \theta_k)}{C(z_1, \ldots, z_n)} .$$

Indem man (19) mit der Verteilung $f(y;\theta_1, \ldots, \theta_k, x_{1o}, \ldots, x_{ro})$

multipliziert, erhält man die gemeinsame Verteilung für $(y,\theta_1,\ldots,\theta_k)$ in Abhängigkeit von $(x_{1o},\ldots,x_{ro};z_1,\ldots,z_n)$. Durch Integration über alle θ_i erhält man schliesslich die gesuchte Verteilung für y:

$$(20)\quad \begin{cases} f^*(y;x_{1o},\ldots,x_{ro}|\,z_1,\ldots,z_n) \\ = \int_{\theta_k}\ldots\int_{\theta_1} f(y;\theta_1,\ldots,\theta_k,x_{1o},\ldots,x_{ro})\,p_B(\theta_1,\ldots,\theta_k|\,z_1,\ldots,z_n)\,d\theta_1\ldots d\theta_k. \end{cases}$$

Für die Anwendung dieses Vorgehens auf einen Regressionsansatz $y_t=a+b(x_t-\bar{x})+\varepsilon_t$ wollen wir das Ergebnis zitieren[1)2)]:

Die Grösse y_t sei $N(a+b(x_t-\bar{x}),\sigma^2)$ verteilt, wobei a,b,σ^2 unbekannt sind. Es liegen Beobachtungen (y_t,x_t), $t=1,\ldots,n$, vor, und ausserdem sei für $a,b,\ln\sigma^2$ je eine Gleichverteilung als a-priori-Verteilung gegeben. Es gelten ferner die Bezeichnungen:

$$\begin{aligned} \hat{a} &= \bar{y} = \frac{1}{n}\sum y_t \\ \hat{b} &= \sum(x_t-\bar{x})(y_t-\bar{y})/\sum(x_t-\bar{x})^2, \quad \bar{x}=\frac{1}{n}\sum x_t \\ S_{xx} &= \sum(x_t-\bar{x})^2 \\ S^2 &= \sum(y_t-\bar{y})^2 - \hat{b}\sum(x_t-\bar{x})(y_t-\bar{y}). \end{aligned}$$

Dann besitzt bei vorgegebenem x_o der Ausdruck

$$(21)\quad \frac{y_t-\hat{a}-\hat{b}(x_o-\bar{x})}{\left\{S^2\left[1+\frac{1}{n}+\frac{(x_o-\bar{x})^2}{S_{xx}}\right]/(n-2)\right\}^{1/2}}$$

eine t - Verteilung mit $(n-2)$ Freiheitsgraden.

Dieses Ergebnis gibt uns die Möglichkeit, die unter gleichen Bedingungen entsprechende Verteilung für K_t bei vorgegebenem $R_{1_M,t-\tau}$ anzugeben.

---

1) Wegen der Herleitung desselben vgl. [11], Teil II, Seite 212-213.

2) Bezüglich der Anwendung der Bayesschen Methode auf allgemeinere lineare Modelle vgl. man [2], [7].

§ 5 DER ENTSCHEIDUNGSTHEORETISCHE ASPEKT BEI DER VORAUSRECHNUNG VON R_{lat_p}

5.1 Vorbemerkung

Wir gehen jetzt davon aus, dass wir aus gegebenen Informationen über die erklärenden Grössen von R_{lat_p} mittels der soeben beschriebenen Verfahren Verteilungen φ_i konstruiert und aus denselben vermittelst (12) auf analytischem oder numerischem Weg eine Wahrscheinlichkeitsverteilung $\omega(R_{lat_p})$ ermittelt haben.

Wir interessieren uns dann dafür, welches bei dieser gegebenen Verteilung die in der Angabe eines einzigen Wertes bestehende optimale Prognose[1] für R_{lat_p} ist. Zur Beantwortung dieser Frage werden wir entscheidungstheoretische Hilfsmittel zu Rate ziehen.

Im Rahmen dieser Untersuchungen werden wir, ebenfalls auf der Basis entscheidungstheoretischer Auffassung, gleichzeitig auch nach einer Lösung des folgenden Problems suchen:

---

1) In diesem Zusammenhang sprechen wir vorzugsweise von Prognose anstatt von Vorausrechnung, da wir die Vorausrechnungs-Phase der Ermittlung von $\omega(R_{lat_p})$ als hinter uns liegend betrachten. Der Begriff Prognose ist hier allerdings nicht in dem auf Seite 1 gebrauchten Sinn zu sehen, sondern im viel engeren Sinn der Vorhersage eines stochastischen Ereignisses oder wie man sagt der stochastischen Prognose (vgl.[17]) bei gegebener Wahrscheinlichkeitsverteilung. In solchem Zusammenhang spricht man von dieser Wahrscheinlichkeitsverteilung als von einer Prognoseverteilung (engl.: predictive density, vgl. hierzu [2]).

Wir hatten oben gesehen, dass uns das Modell (9,11) zwar im Prinzip den Zusammenhang zwischen $\omega(R_{lat_p})$ und den Verteilungen φ_i der erklärenden Variablen schildert, dass es aber in konkreten Fällen wegen der Kompliziertheit dieses Zusammenhangs nicht möglich oder zumindest nicht rationell sein wird, die Verteilung $\omega(R_{lat_p})$ aus gegebenen Verteilungen φ_i zu bestimmen. Es soll jetzt - wie oben bereits angedeutet wurde - herausgefunden werden, auf welche Parameter der Verteilung $\omega(R_{lat_p})$ wir unsere Betrachtung beschränken können, ohne für die Situation relevante Information zu verlieren.

Diese Beschränkung muss sich natürlich an einem übergeordneten Kriterium orientieren, welches wir gerade in einer Erweiterung unserer Perspektive auf die Entscheidungstheorie finden können.

5.2 Die Prognose von R_{lat_p} als Entscheidungsproblem - die Bayessche Prognose und das Bayessche Risiko

Wir beginnen damit, das Prognoseproblem von R_{lat_p} als Entscheidungsproblem zu formulieren.

Bei Anlehnung an die geläufigen Grundbegriffe der Entscheidungstheorie[1] stellt die unbekannte zukünftige Grösse R_{lat_p} den unbekannten Zustand der Natur dar; der Zustandsraum, den wir mit Ω bezeichnen, wird repräsentiert durch die Gesamtheit der möglichen Werte von R_{lat_p}. Die möglichen Prognosen für R_{lat_p}, welche wir mit $\hat{R}_{lat_p}$ bezeichnen, stellen

1) Wegen des hier zugrundegelegten entscheidungstheoretischen Konzepts und insbesondere wegen der einzelnen Begriffe vergleiche man zum Beispiel [15],[19],[24],[25],[27],[32].

die möglichen Aktionen dar, von welchen angesichts der herrschenden Ungewissheit eine auszuwählen ist; der mit $\hat{\Omega}$ bezeichnete Aktionsraum wird gebildet durch die Gesamtheit der möglichen Prognosen. Da alle möglichen Ereignisse für R_{lat_p} gleichzeitig auch mögliche Prognosen sind, besteht die Identität $\hat{\Omega} = \Omega$.

Mit der Wahl einer Prognose $\hat{R}$ [1] verbinden sich gewisse Konsequenzen: wenn die Prognose $\hat{R}$ getroffen wird und sich in Wirklichkeit der Wert R einstellt, dann realisiert sich ein reellwertiger nichtnegativer Verlust $L(\hat{R},R)$, welchen man sich üblicherweise so normiert vorstellt, dass er für $\hat{R} = R$ gleich Null ist. Die auf diese Weise über dem Produktraum $\hat{\Omega} \times \Omega$ definierte Funktion L ist die Verlustfunktion des Prognoseproblems.

Das Prognoseproblem lautet dann folgendermassen: Welches ist unter dem Aspekt der bei Fehlprognosen zu erleidenden Verluste und in Anbetracht der gegebenen Verteilung $\omega(R)$ über Ω die optimale Prognose für R?

Diese Entscheidung vertrauen wir dem Bayesschen Entscheidungskriterium[2] an, welches uns empfiehlt, diejenige Prognose zu wählen, für welche der im Hinblick auf die Verteilung $\omega(R)$[3] zu erwartende Verlust minimal ist. Diese optimale Prognose bezeichnen wir in Erinnerung an das angewandte Kriterium mit $\hat{R}_B$ (B = Bayes).

In mathematischer Formulierung lautet dieses Kriterium folgendermassen:

---

1) Der Einfachheit halber schreiben wir, solange kein Missverständnis möglich ist, nur R anstatt R_{lat_p}.

2) Angesichts der als bekannt angenommenen Verteilung $\omega(R)$ kommt als Entscheidungskriterium nur das als Inbegriff rationalen Verhaltens geltende Bayes-Kriterium in Betracht (vgl. hierzu [19], Seite 191, sowie [28]).

3) Wir fassen R als kontinuierliche Grösse und $\omega(R)$ dementsprechend als Dichtefunktion auf.

Es ist diejenige Prognose $\hat{R}_B$ optimal, für welche die Relation

(22) $\int_{\Omega} L(\hat{R}_B,R)\,\omega(R)dR = \inf_{\hat{R}\in\hat{\Omega}} \int_{\Omega} L(\hat{R},R)\,\omega(R)dR$

erfüllt ist.

Die Prognose $\hat{R}_B$ ist also im Sinne des Bayes-Kriteriums optimal -wir bezeichnen sie deshalb auch als Bayessche Prognose. Den mit einer Prognose $\hat{R}$ verbundenen Erwartungsverlust $r(\hat{R},\omega;L) = \int_{\Omega} L(\hat{R},R)\,\omega(R)dR$ bezeichnen wir als Prognoserisiko. Das Minimum dieses Prognoserisikos, welches laut Definition durch $\hat{R}_B$ erreicht wird und welches durch (22) gegeben ist, bezeichnen wir als Bayessches Prognoserisiko oder kurz als Bayessches Risiko.

Das Bayessche Risiko stellt in gewissem Sinn einen Maßstab dar zur Beurteilung des Ausmaßes oder der Tragweite der um R herrschenden Ungewißheit hinsichtlich der Konsequenzen bei falschen Prognosen. Es mißt dabei gewissermaßen auch den Wert der in $\omega(R)$ enthaltenen Information über R. Dieser wichtigen Eigenschaft des Bayesschen Risikos werden wir uns an späterer Stelle noch bedienen.

Wenn übrigens die Verlustfunktion von der speziellen Form $L(\hat{R},R) = (\hat{R}-R)^2$ ist, dann ist das mit einer Prognose $\hat{R}$ verbundene Risiko $\int_{\Omega}(\hat{R}-R)^2\,\omega(R)dR$ bekannt als der mittlere quadratische Fehler der Prognose $\hat{R}$.[1)]

---

1) Dieser mittlere quadratische Fehler mißt ebenfalls das Ausmaß der Ungewißheit einer Prognose $\hat{R}$ und im Fall einer Bayesschen Prognose (welche in diesem Fall , vgl.5.3.a, gleich dem Erwartungswert von R ist) das Ausmaß der herrschenden Ungewißheit überhaupt. Obwohl diese Grösse ebenfalls eine Verlustfunktion impliziert, erlaubt sie wegen der Unveränderlichkeit derselben dennoch in bezug auf die Konsequenzen bei Fehlprognosen keine spezielle Einschätzung einer Situation. Sie beurteilt also die Ungewißheit nicht wie das Risikomaß im Hinblick auf Konsequenzen. Diese Grösse ist also dann als Ungewißheitsmaß angebracht, wenn keine Verlustfunktion vorliegt oder wenn diesbezüglich Indifferenz herrscht.

Aus der Sicht des obigen entscheidungstheoretischen Ansatzes sind die Bayessche Prognose und das Bayessche Risiko die einzigen im Zusammenhang mit einem Prognoseproblem relevanten Grössen. Für die Informationsbeschaffung über $\omega(R)$ bedeutet das also, dass nur diejenigen Parameter von $\omega(R)$ von Bedeutung sind, welche die Bayessche Prognose und das Bayessche Risiko bestimmen[1].

5.3 Die optimale Prognose für R_{lat_p} bei speziellen Verlustfunktionen

Um zu zeigen, durch welche Parameter von $\omega(R)$ die Bayessche Prognose und das Bayessche Risiko bestimmt sein können, wollen wir jetzt einige spezielle Verlustfunktionen betrachten.

a) Bei der quadratischen Verlustfunktion $L(\hat{R},R)=k(\hat{R}-R)^2$ $(k>o)$ ist die Bayessche Prognose durch den Erwartungswert[2] $E(R)$ gegeben und das Bayessche Risiko durch $k\sigma^2(R)$, wobei $\sigma^2(R)$ die Streuung von R bezeichnet.

Beweis: Zum Beweis haben wir zu zeigen, dass $E(R)$ derjenige Wert für $\hat{R}$ ist, für welchen die 1.Ableitung des Prognoserisikos $k\int_{\Omega}(\hat{R}-R)^2\omega(R)dR$ nach $\hat{R}$ gleich Null ist[3].

$$\frac{\partial}{\partial\hat{R}}\, k\int_{\Omega}(\hat{R}-R)^2\omega(R)dR = \frac{\partial}{\partial\hat{R}}\, k\Big[\hat{R}^2\int_{\Omega}\omega(R)dR-2\hat{R}\int_{\Omega}R\omega(R)dR+\int_{\Omega}R^2\omega(R)dR\Big]$$

$$= k[2\hat{R} - 2E(R)].$$

---

1) In diesem Sinn sind also Verteilungen, welche dieselbe Bayessche Prognose und dasselbe Bayessche Risiko liefern, als äquivalent zu betrachten.

2) Dieses Ergebnis wirft nochmals ein gewisses Licht auf das in § 3 entwickelte Modell, bei welchem wir uns auf die Vorausrechnung der Erwartungswerte beschränkt hatten. Mit jenem Modell erstellte Vorausrechnungen sind also aus entscheidungstheoretischer Sicht bei einer quadratischen Verlustfunktion optimal.

3) Die nachher folgenden Beweise werden nach demselben Prinzip geführt werden. Es ist aber zu beachten, dass diese Beweise erst vollständig sind, wenn man gezeigt hat, dass an der betreffenden Stelle die 2.Ableitung nach $\hat{R}$ positiv ist (vgl. hierzu [23], Band 1, Seite 322 ff). Wir werden dies jedoch nicht im einzelnen ausführen.

Der letzte Ausdruck ist genau dann gleich Null, wenn $\hat{R}=E(R)$. Durch Einsetzen von $E(R)$ in den Ausdruck $k\int_{\Omega}(\hat{R}-R)^2\,\omega(R)dR$ erhält man auf Grund der Definition der Streuung die Grösse $k\sigma^2(R)$.

b) Bei der linearen Verlustfunktion $L(\hat{R},R)=k|\hat{R}-R|\,(k>0)$ ist die Bayessche Prognose durch den Median $M(R)$ gegeben und das Bayessche Risiko durch $k\int_{\Omega}|M(R)-R|\,\omega(R)dR$ (der Integralausdruck stellt die mittlere absolute Abweichung zwischen R und M(R) dar).

Beweis :

$$\frac{\partial}{\partial\hat{R}}\,k\int_{\Omega}|\hat{R}-R|\,\omega(R)dR$$

$$=\frac{\partial}{\partial\hat{R}}\,k\Big[\int_{-\infty}^{\hat{R}}(\hat{R}-R)\,\omega(R)dR+\int_{\hat{R}}^{\infty}(R-\hat{R})\,\omega(R)dR\Big]$$

$$=\frac{\partial}{\partial\hat{R}}\Big\{k\Big[\hat{R}\int_{-\infty}^{\hat{R}}\omega(R)dR-\int_{-\infty}^{\hat{R}}R\,\omega(R)dR\Big]+k\Big[\int_{\hat{R}}^{\infty}R\omega(R)dR-\hat{R}\int_{\hat{R}}^{\infty}\omega(R)dR\Big]\Big\}$$

$$=k\Big[\int_{-\infty}^{\hat{R}}\omega(R)dR+\hat{R}\,\omega(\hat{R})-\hat{R}\,\omega(\hat{R})-\hat{R}\,\omega(\hat{R})-\int_{\hat{R}}^{\infty}\omega(R)dR+\hat{R}\,\omega(\hat{R})\Big]$$

$$=k\Big[\int_{-\infty}^{\hat{R}}\omega(R)dR-\int_{\hat{R}}^{\infty}\omega(R)dR\Big]\,.$$

Der letzte Ausdruck ist genau dann gleich Null, wenn $\int_{-\infty}^{\hat{R}}\omega(R)dR=\frac{1}{2}$ ist, d.h. wenn $\hat{R}$ der Median von R ist. Durch Einsetzen des Medians in $k\int_{\Omega}|\hat{R}-R|\,\omega(R)dR$ erhält man das Bayessche Risiko.

c) Bei der asymmetrischen quadratischen Verlustfunktion

$$L(\hat{R},R)=\begin{cases}k_1(\hat{R}-R)^2 \text{ für } \hat{R}\geq R\\ k_2(\hat{R}-R)^2 \text{ für } \hat{R}<R\end{cases}\qquad k_1,k_2>0$$

ist die Bayessche Prognose derjenige Punkt $\hat{R}_B$

der Verteilung $\omega(R)$, für welchen die folgende Relation gilt:

$$\frac{k_1}{k_2} = \frac{\int_{\hat{R}_B}^{\infty} (R-\hat{R}_B)\,\omega(R)dR}{\int_{-\infty}^{\hat{R}_B} (\hat{R}_B-R)\,\omega(R)dR} \quad .$$

$\hat{R}_B$ ist aus physikalischer Sicht derjenige Unterstützungspunkt einer über eine Achse R nach $\omega(R)$ verteilten Masse, bei welchem die links- und rechtswirkenden Drehmomente im Verhältnis k_1/k_2 stehen. Ein einfacher Ausdruck für das Bayessche Risiko kann nicht angegeben werden.

Beweis:

$$\frac{\partial}{\partial \hat{R}}\left[k_1 \int_{-\infty}^{\hat{R}} (\hat{R}-R)^2\omega(R)dR + k_2 \int_{\hat{R}}^{\infty} (\hat{R}-R)^2\omega(R)dR\right]$$

$$= \frac{\partial}{\partial \hat{R}}\left[k_1\hat{R}^2 \int_{-\infty}^{\hat{R}} \omega(R)dR - 2k_1\hat{R}\int_{-\infty}^{\hat{R}} R\omega(R)dR + k_1\int_{-\infty}^{\hat{R}} R^2\omega(R)dR\right.$$

$$\left. +k_2\hat{R}^2 \int_{\hat{R}}^{\infty} \omega(R)dR - 2k_2\hat{R}\int_{\hat{R}}^{\infty} R\omega(R)dR + k_2\int_{\hat{R}}^{\infty} R^2\omega(R)dR\right]$$

$$= 2k_1\hat{R}\int_{-\infty}^{\hat{R}}\omega(R)dR + k_1\hat{R}^2\omega(\hat{R}) - 2k_1\int_{-\infty}^{\hat{R}} R\omega(R)dR - 2k_1\hat{R}^2\omega(\hat{R}) + k_1\hat{R}^2\omega(\hat{R})$$

$$+2k_2\hat{R}\int_{\hat{R}}^{\infty}\omega(R)dR - k_2\hat{R}^2\omega(\hat{R}) - 2k_2\int_{\hat{R}}^{\infty} R\omega(R)dR + 2k_2\hat{R}^2\omega(\hat{R}) - k_2\hat{R}^2\omega(\hat{R})$$

$$= 2k_1\int_{-\infty}^{\hat{R}} (\hat{R}-R)\omega(R)dR - 2k_2\int_{\hat{R}}^{\infty} (R-\hat{R})\omega(R)dR \quad .$$

Der letzte Ausdruck ist genau dann gleich Null wenn

$$\frac{k_1}{k_2} = \frac{\int_{\hat{R}}^{\infty}(R-\hat{R})\,\omega(R)dR}{\int_{-\infty}^{\hat{R}}(\hat{R}-R)\,\omega(R)dR} \quad .$$

d) Bei der asymmetrischen linearen Verlustfunktion[1]

$$L(\hat{R},R)=\begin{cases} k_1(\hat{R}-R) & \text{für } \hat{R}\geq R \\ k_2(R-\hat{R}) & \text{für } \hat{R}<R \end{cases} \qquad k_1,k_2>0$$

ist die Bayessche Prognose gegeben durch das $\frac{k_2}{k_1+k_2}$-Quantil der Verteilung $\omega(R)$, d.h. also durch den Punkt $\hat{R}_B$, für welchen gilt $pr(R<\hat{R}_B)=\frac{k_2}{k_1+k_2}$.

Beweis :

$$\frac{\partial}{\partial\hat{R}}\Big[k_1\int_{-\infty}^{\hat{R}}(\hat{R}-R)\omega(R)dR+k_2\int_{\hat{R}}^{\infty}(R-\hat{R})\omega(R)dR\Big]$$

$$=\frac{\partial}{\partial\hat{R}}\Big[k_1\hat{R}\int_{-\infty}^{\hat{R}}\omega(R)dR-k_1\int_{-\infty}^{\hat{R}}R\omega(R)dR+k_2\int_{\hat{R}}^{\infty}R\omega(R)dR-k_2\hat{R}\int_{\hat{R}}^{\infty}\omega(R)dR\Big]$$

$$=k_1\int_{-\infty}^{\hat{R}}\omega(R)dR+k_1\hat{R}\omega(\hat{R})-k_1\hat{R}\omega(\hat{R})-k_2\hat{R}\omega(\hat{R})-k_2\int_{\hat{R}}^{\infty}\omega(R)dR+k_2\hat{R}\omega(\hat{R})$$

$$=k_1\int_{-\infty}^{\hat{R}}\omega(R)dR-k_2\int_{\hat{R}}^{\infty}\omega(R)dR$$

$$=k_1\int_{-\infty}^{\hat{R}}\omega(R)dR-k_2\Big[1-\int_{-\infty}^{\hat{R}}\omega(R)dR\Big]=(k_1+k_2)\int_{-\infty}^{\hat{R}}\omega(R)dR-k_2.$$

Der letzte Ausdruck ist genau dann gleich Null, wenn

$$\int_{-\infty}^{\hat{R}}\omega(R)dR=k_2/(k_1+k_2) \quad \text{ist.}$$

e) Bei der Verlustfunktion[2]

$$L(\hat{R},R)=\begin{cases} 0 & \text{für } |\hat{R}-R|\leq c \\ k & \text{für } |\hat{R}-R|>c \end{cases} \qquad k>0,\ c\geq 0,$$

ist die Bayessche Prognose gleich der Mitte des Modalintervalls der Länge 2 c der Verteilung $\omega(R)$. Für c=o ergibt sich als Bayessche Prognose der Modalwert von $\omega(R)$.

---

1) Diese Verlustfunktion wird auch in [25], Seite 196 untersucht.

2) Diese Verlustfunktion wird in [32], Seite 23 erwähnt.

Beweis : $\frac{\partial}{\partial \hat{R}} \int_{-\infty}^{\hat{R}-c} k\,\omega(R)dR + \int_{\hat{R}+c}^{\infty} k\omega(R)dR = k\omega(\hat{R}-c)-k\omega(\hat{R}+c).$

Dieser Ausdruck ist gleich Null für $\omega(\hat{R}-c)=\omega(\hat{R}+c)$, wodurch das Modalintervall der Länge 2c festgelegt ist[1].

f) Zum Abschluss wollen wir noch eine Verlustfunktion betrachten, bei welcher die Verluste von den relativen Abweichungen zwischen $\hat{R}$ und R abhängen.
Bei der Verlustfunktion $L(\hat{R},R) = k\left(\frac{\hat{R}-R}{\hat{R}}\right)^2$, $k>o$, $\hat{R}\neq o$, ist die Bayessche Prognose gegeben durch

$$\hat{R}_B = \frac{\sigma^2(R)+E^2(R)}{E(R)} = E(R)[V^2(R)+1]\,,^{2)}$$

wobei $V(R)=\sigma(R)/E(R)$ der Variationskoeffizient von R ist.

Das Bayessche Risiko ist gegeben durch den Ausdruck

$$\frac{k}{\hat{R}_B^2}\left[\sigma^2(R)+(E(R)-\hat{R}_B)^2\right] = k\cdot\frac{V^2(R)}{V^2(R)+1}\,.$$

Wenn V(R) klein ist gegenüber 1, dann ist $\hat{R}_B$ annähernd gleich E(R) und das Bayessche Risiko annähernd gleich $kV^2(R)$.

Beweis : $\frac{\partial}{\partial \hat{R}} k\int_{\Omega}\left(\frac{\hat{R}-R}{\hat{R}}\right)^2\omega(R)dR = \frac{2kE(R)}{\hat{R}^2} - \frac{2k}{\hat{R}^3}\left[\sigma^2(R)+E^2(R)\right].$

Dieser Ausdruck ist gleich Null für $\hat{R} = \frac{\sigma^2(R)+E^2(R)}{E(R)}$.

Durch Einsetzen dieses Ausdrucks in $\frac{k}{\hat{R}^2}\int_{\Omega}(\hat{R}-R)^2\omega(R)dR$ erhält man nach einer einfachen Umformung unter Berücksichtigung von $V(R)=\sigma(R)/E(R)$ den Ausdruck $E(R)[V^2(R)+1]$.

Die Approximationen für kleine Werte von V(R) ergeben sich durch Vernachlässigung von $V^2(R)$.

Ansätze mit Verlustfunktionen der Art $L(\hat{R},R) = k|\hat{R}-R|^{\alpha}$, $o<\alpha<1$, $k>o$, bei welchen das Wachstum der Verluste degressiv ist, können

---

1) Es wird hier vorausgesetzt, dass $\omega(R)$ nur einen Gipfel hat.

2) Bei dieser Verlustfunktion darf E(R) nicht Null sein.

zwar von praktischer Bedeutung sein, wir wollen hier jedoch nicht näher auf sie eingehen, da sie nur approximativ gelöst werden können oder numerisch zu behandeln sind. Die obigen Ergebnisse zeigen, wie sich das optimale Verhalten bei einer Prognose in Abhängigkeit von der Verlustfunktion ändert. Wie ausgeprägt diese Abhängigkeit ist, wird andererseits auch von der Verteilung $\omega(R)$ mitbestimmt, wobei insbesondere ihr Symmetriegrad eine grosse Rolle spielen kann. Wie präzise die Verlustfunktion bekannt sein muss, bestimmt sich aus ihrem Zusammenwirken mit der Verteilung $\omega(R)$. Wir wollen auf Einzelheiten hierüber nicht eingehen, auch nicht auf eine Diskussion der praktischen Angemessenheit der oben aufgeführten Typen von Verlustfunktionen.

5.4 Die praktische Relevanz des Begriffs der Verlustfunktion beim Problem der Vorausrechnung von Rinderprozessen.

Es ist angebracht, den Begriff der Verlustfunktion bezüglich der praktischen Relevanz bei unserem Problem der Vorausrechnung von Rinderprozessen zu untersuchen, um letztlich festzustellen, ob der obige entscheidungstheoretische Ansatz überhaupt dem Wesen des Problems angepasst ist.

Um dies zu beurteilen, vergegenwärtigen wir uns die möglichen Verwendungszwecke für Vorausrechnungen von Rinderprozessen. Auf der Basis solcher Vorausrechnungen werden z.B. Fleisch- und Milchversorgungsmaßnahmen getroffen, Importe bzw. Exporte vereinbart, Interventionspreise festgelegt, Subventionsmittel bereitgestellt, u.s.w.. In allen diesen Fällen handelt es sich um Entscheidungen, welche, wenn sie sich als Fehlentscheidungen herausstellen, materielle Verluste nach sich ziehen. Diese Entscheidungen werden aber umso besser sein, ihre Konsequenzen also umso weniger gravierend, je treffender die Vorausrechnungen sind; oder umgekehrt, je ungenauer die Vorausrechnungen sind, desto schlechtere Entscheidungen werden auf ihrer Grundlage getroffen und desto schwerwiegender werden die Folgen sein.

Vorausrechnungen von Rinderprozessen liefern also wichtige Informationen, welche als Orientierungsgrundlage bei der Wahl von Entscheidungen in verlustbedrohten Situationen auf dem Rindersektor und in benachbarten Sektoren eine maßgebliche Rolle spielen. Abweichungen der Vorausrechnungen von der Realität bedingen also materielle Verluste, so daß der obige entscheidungstheoretische Ansatz auf der Basis einer Verlustfunktion als dem Problem angemessen gelten kann.

Wir wollen an dieser Stelle noch zeigen, wie die Verlustfunktion $L(\hat{R},R)$ des Vorausrechnungsproblems in Zusammenhang gebracht werden kann mit der Verlustfunktion $L(a,R)$ eines allgemeinen Entscheidungsproblems, bei welchem angesichts des unbekannten Zustands der Natur R eine Entscheidung zwischen gewissen Aktionen a aus einer Menge A zu treffen ist (diese Aktionen können dabei zum Beispiel in agrarpolitischen Massnahmen oder Vereinbarungen bestehen, wie wir sie oben erwähnt haben).

Dem Entscheidenden unterstellen wir die folgende Strategie[1)2)]: Wenn er über eine Vorausrechnung $\hat{R}$ für R verfügt, dann orientiert er seine Wahl in A ausschliesslich an $\hat{R}$ und wählt diejenige Aktion $a_{\hat{R}}$, für welche gilt:

$$(23) \qquad L(a_{\hat{R}},\hat{R}) = \inf_{a \in A} L(a,\hat{R})$$ [3)]

Bei dieser Aktion entsteht ihm also beim Eintreten des wahren Ereignisses R der durch $\hat{R}$ und R bestimmte Verlust $L(a_{\hat{R}},R)$. Die durch $L^*(\hat{R},R) = L(a_{\hat{R}},R)$ gegebene Verlustfunktion kann man als Verlustfunktion für das Vorausrechnungsproblem deuten.

Wenn man von $L^*(\hat{R},R)$ noch den Verlust $L(a_R,R)$ subtrahiert, welchen der Entscheidende auch bei exakter Kenntnis von R und entsprechend idealer Entscheidung a_R hätte hinnehmen müssen, so besitzt die entstehende Verlustfunktion $L^{**}(\hat{R},R)=L(a_{\hat{R}},R)-L(a_R,R)$ offenbar auch die Normeigenschaft, für $\hat{R} = R$ gleich Null und sonst nichtnegativ zu sein.

---

1) Wegen des Strategiebegriffs im allgemeinen vgl. z.B.[32].

2) Zu dieser speziellen Strategie vgl.[6].

3) Dieses Verhalten dürfte übrigens in der Praxis gar nicht unüblich sein. Wegen einer Deutung eines solchen Verhaltens auf der Basis der Likelihood-Funktion vgl.[5].

5.5 Die den weiteren Untersuchungen zugrunde gelegte Verlustfunktion

Für weitere, auf entscheidungstheoretischer Grundlage beruhende Untersuchungen über das Vorausrechnungsproblem von R_{lat_p} legen wir die Verlustfunktionen

$$L_1(\hat{R},R) = k(\hat{R}-R)^2$$

und

$$L_2(\hat{R},R) = k\left(\frac{\hat{R}-R}{\hat{R}}\right)^2$$

zugrunde.

Mit ersterer gedenken wir diejenigen Situationen zu erfassen, bei welchen der Verlust von der absoluten Abweichung zwischen Vorausrechnung und wirklichem Wert abhängt, mit der zweiten können wir den Fällen Rechnung tragen, bei welchen der Verlust von der relativen Abweichung zwischen Vorausrechnung und wirklichem Wert bestimmt wird.

Diese Vereinbarung über die Verlustfunktion treffen wir in erster Linie wegen der sich mit diesen beiden Verlustfunktionen verbindenden mathematischen Vereinfachungen. Bei anderen Ansätzen über die Verlustfunktion wären nämlich die im folgenden zu behandelnden Probleme nicht analytisch und vermutlich auch nicht mit numerischen Verfahren lösbar.

Mit dieser Vereinbarung ist laut 5.3.a und 5.3.f die (optimale) Bayessche Prognose durch $E(R)$ bzw. durch $E(R)[V^2(R)+1]$ und das damit verbundene Bayessche Risiko durch $k\cdot\sigma^2(R)$ bzw. durch $kV^2(R)/[V^2(R)+1]$ gegeben. Mittelwert und Streuung bzw. Variationskoeffizient von R_{lat_p} sind also aus unserer entscheidungstheoretischen Perspektive in Verbindung mit der getroffenenen Vereinbarung über die Verlustfunktion die einzigen im Zusammenhang mit der Vorausrechnung von R_{lat_p} relevanten Parameter der Verteilung $\omega(R_{lat_p})$. Wir können uns also im Rahmen des obigen stochastischen Vorausrechnungsmodells auf die Bestimmung dieser Parameter beschränken.

Als nächstes werden wir uns der Aufgabe zuwenden, den Mittelwert, die Streuung und den Variationskoeffizienten von R_{lat_p} aus den Mittelwerten, Streuungen und Variationskoeffizienten der erklärenden Grössen von R_{lat_p} zu bestimmen.

§6 BESTIMMUNG DES MITTELWERTES, DER STREUUNG UND DES VARIATIONSKOEFFIZIENTEN VON R_{lat_p}

Wir werden jetzt den Mittelwert, die Streuung und den Variationskoeffizienten von R_{lat_p} durch die Mittelwerte, Streuungen und Variationskoeffizienten der erklärenden Grössen $R_{1,a-t_p+t_o,t_o}$, $p_{lat_pt_o}$, K_{t_p-a}, q_{1a} und γ_1 ausdrücken.
Den Mittelwert von R_{lat_p} bezeichnen wir mit $E(R_{lat_p})$, die Streuung mit $\sigma^2(R_{lat_p})$ und den Variationskoeffizienten, welcher definiert ist als $\sigma(R)/E(R)$, mit $V(R_{lat_p})$.

Aus (9) und (10) ergeben sich zunächst die folgenden bedingten Mittelwerte und Streuungen[1]:

$$(24)\quad \begin{cases} E(R_{lat_p} \mid R_{1,a-t_p+t_o,t_o}, p_{lat_pt_o}) = R_{1,a-t_p+t_o,t_o} \cdot p_{lat_pt_o}, & \text{für } a > t_p-t_o \\ E(R_{lat_p} \mid R_{1,o,t_p-a}, q_{1a}) = R_{1,o,t_p-a} \cdot q_{1a}, & \text{für } a < t_p-t_o, \end{cases}$$

wobei R_{1,o,t_p-a} der binomialverteilten Grösse $\sum_{j=1}^{K_{t_p-a}} z_j^{**}$ aus (9) entspricht.

$$(25)\quad \begin{cases} \sigma^2(R_{lat_p} \mid R_{1,a-t_p+t_o,t_o}, p_{lat_pt_o}) \\ = R_{1,a-t_p+t_o,t_o} \cdot p_{lat_pt_o}(1-p_{lat_pt_o}) & \text{für } a > t_p-t_o \\ \sigma^2(R_{lat_p} \mid R_{1,o,t_p-a}, q_{1a}) \\ = R_{1,o,t_p-a} \cdot q_{1a}(1-q_{1a}) & \text{für } a < t_p-t_o. \end{cases}$$

1) Es liegen dabei die Formeln des Mittelwertes und der Streuung einer B(n,p)-binomialverteilten Grösse x zugrunde:
$E(x) = n \cdot p$ und $\sigma^2(x) = np(1-p)$ (vgl. hierzu [9], Seite 116).

Um aus diesen bedingten Grössen in Verbindung mit den Mittelwerten und Streuungen der Grössen $R_{1,a-t_p+t_o,t_o}$, $P_{lat_p t_o}$, K_{t_p-a}, q_{la} und γ_1 die marginalen Mittelwerte und Streuungen $E(R_{lat_p})$ und $\sigma^2(R_{lat_p})$ zu berechnen, benötigen wir die Ergebnisse des folgenden Satzes über die Beziehungen zwischen bedingten und marginalen Mittelwerten und Streuungen.

Satz: Gegeben seien stochastische Veränderliche $y, x_1, \ldots, x_k$. Wenn $E(y)$ den Erwartungswert von y, $\sigma^2(y)$ die Streuung von y, $E(y|x_1,\ldots,x_k)$ den bedingten Erwartungswert von y bei gegebenem $(x_1,\ldots,x_k)$[1], $\sigma^2(y|x_1,\ldots,x_k)$ die bedingte Streuung von y bei gegebenem $(x_1,\ldots,x_k)$, und $E_{x_1,\ldots,x_k}$ die Bildung des Erwartungswertes einer Funktion von $(x_1,\ldots,x_k)$ bezüglich $(x_1,\ldots,x_k)$ bezeichnet, dann gelten unter Voraussetzung der Existenz der genannten Grössen die beiden folgenden Relationen:

(26) $$E(y) = E_{x_1,\ldots,x_k} E(y|x_1,\ldots,x_k)$$

(27) $$\sigma^2(y) = E_{x_1,\ldots,x_k}[\sigma^2(y|x_1,\ldots,x_k)+E^2(y|x_1,\ldots,x_k)] - E^2(y).$$

Beweis: ad (26).

Es sei $f(y)$ die Dichte der Verteilung von y, $f_1(y|x_1,\ldots,x_k)$ die Dichte von y unter der Bedingung $(x_1,\ldots,x_k)$ und $f_2(x_1,\ldots,x_k)$ die Dichte von $(x_1,\ldots,x_k)$. Dann gilt:

---

1) Der Ausdruck $E(y|x_1,\ldots,x_k)$ ist als Regression von y bezüglich $(x_1,\ldots,x_k)$ bekannt (vgl. hierzu [9], Seite 82 ff).

$$E(y) = \int_y y\, f(y)dy$$

$$= \int_y \int_{x_k} \dots \int_{x_1} y\, f_1(y|x_1,\dots,x_k) f_2(x_1,\dots,x_k)dx_1\dots dx_k dy$$

$$= \int_{x_k} \dots \int_{x_1} E(y|x_1,\dots,x_k) f_2(x_1,\dots,x_k)dx_1\dots dx_k$$

$$= E_{x_1,\dots,x_k} E(y|x_1,\dots,x_k).$$

ad (27).

Bekanntlich gilt die Zerlegung

$$\sigma^2(y) = E(y^2) - E^2(y).$$

Da analog zu (26) auch gilt

$$E(y^2) = E_{x_1,\dots,x_k} E(y^2|x_1,\dots,x_k),$$

erhält man

$$\sigma^2(y) = E_{x_1,\dots,x_k} E(y^2|x_1,\dots,x_k) - E^2(y).$$

Da ferner gilt

$$E(y^2|x_1,\dots,x_k) = \sigma^2(y|x_1,\dots,x_k) + E^2(y|x_1,\dots,x_k),$$

ergibt sich schliesslich

$$\sigma^2(y) = E_{x_1,\dots,x_k}[\sigma^2(y|x_1,\dots,x_k)+E^2(y|x_1,\dots,x_k)] - E^2(y).$$

Mit diesem Ergebnis können wir fortfahren in der Ermittlung von $E(R_{lat_p})$ und $\sigma^2(R_{lat_p})$. Zum Zwecke einer übersichtlicheren Darstellung führen wir unter Vernachlässigung der Indizes[1] die folgenden abkürzenden Bezeichnungen ein:

---

1) D.h. dieselben denken wir uns vorübergehend konstant gehalten.

$$
(28)\quad \begin{cases}
E(R_{l,a-t_p+t_o,t_o}) = m_1 & \sigma^2(R_{l,a-t_p+t_o,t_o}) = \sigma_1^2 \\
E(p_{lat_p t_o}) = m_2 & \sigma^2(p_{lat_p t_o}) = \sigma_2^2 \\
E(K_{t_p-a}) = m_3 & \sigma^2(K_{t_p-a}) = \sigma_3^2 \\
E(q_{la}) = m_4 & \sigma^2(q_{la}) = \sigma_4^2 \\
E(\gamma_l) = m_5 & \sigma^2(\gamma_l) = \sigma_5^2 \\
E(R_{l,o,t_p-a}) = m_6 & \sigma^2(R_{l,o,t_p-a}) = \sigma_6^2 .
\end{cases}
$$

Unter Anwendung von (26) in Verbindung mit (24) erhalten wir

$$
(29)\quad \begin{cases}
E(R_{lat_p}) = m_1 m_2 & \text{für } a \geq t_p - t_o \\
E(R_{lat_p}) = m_6 m_4 & \text{für } a < t_p - t_o .
\end{cases}
$$

Unter Anwendung von (27) in Verbindung mit (25) und Berücksichtigung von (29) erhalten wir

$$
(30)\quad \begin{cases}
\sigma^2(R_{lat_p}) & \\
= E_{R,p}\left[R_{l,a-t_p+t_o,t_o} p_{lat_p t_o}(1-p_{lat_p t_o}) + R^2_{l,a-t_p+t_o,t_o} p^2_{lat_p t_o}\right] - m_1^2 m_2^2 & \\
= m_1 m_2 - m_1^2 m_2^2 + (\sigma_2^2 + m_2^2)(\sigma_1^2 + m_1^2 - m_1) & \text{für } a \geq t_p - t_o \\
\sigma^2(R_{lat_p}) & \\
= E_{R,q}\left[R_{l,o,t_p-a} q_{la}(1-q_{la}) + R^2_{l,o,t_p-a} q^2_{la}\right] - m_6^2 m_4^2 & \\
= m_4 m_6 - m_4^2 m_6^2 + (\sigma_4^2 + m_4^2)(\sigma_6^2 + m_6^2 - m_6) & \text{für } a < t_p - t_o .
\end{cases}
$$

Jetzt werden wir noch m_6 und σ_6^2 durch die Mittelwerte und Streuungen von K_{t_p-a} und γ_1 ausdrücken.

Analog zu (24) bzw. (25) gelten im Hinblick auf (9) die folgenden Relationen:

$$(31) \quad E(R_{1,o,t_p-a} | K_{t_p-a}, \gamma_1) = \gamma_1 \cdot K_{t_p-a}$$

$$(32) \quad \sigma^2(R_{1,o,t_p-a} | K_{t_p-a}, \gamma_1) = K_{t_p-a} \cdot \gamma_1(1-\gamma_1).$$

Daraus folgt unter Berücksichtigung von (26) und (27):

$$(33) \quad E(R_{1,o,t_p-a}) = m_6 = m_5 m_3$$

$$(34) \quad \sigma^2(R_{1,o,t_p-a}) = \sigma_6^2$$
$$= E_{K,\gamma}[K_{t_p-a} \cdot \gamma_1(1-\gamma_1) + \gamma_1^2 \cdot K_{t_p-a}^2] - m_5^2 m_3^2$$
$$= m_5 m_3 - m_5^2 m_3^2 + (\sigma_5^2 + m_5^2)(\sigma_3^2 + m_3^2 - m_3).$$

Durch Einsetzen von (33) in (29) und von (34) in (30) erhalten wir schliesslich als Endergebnis:

$$(35) \begin{cases} E(R_{lat_p}) = m_1 m_2 & \text{für } a \geq t_p - t_o \\ E(R_{lat_p}) = m_3 m_4 m_5 & \text{für } a < t_p - t_o \end{cases}$$

$$(36) \begin{cases} \sigma^2(R_{lat_p}) = m_1 m_2 - m_1^2 m_2^2 + (\sigma_1^2+m_1^2-m_1)(\sigma_2^2+m_2^2) & a \geq t_p - t_o \\ \sigma^2(R_{lat_p}) = m_3 m_4 m_5 - m_3^2 m_4^2 m_5^2 + (\sigma_3^2+m_3^2-m_3)(\sigma_4^2+m_4^2)(\sigma_5^2+m_5^2) & a < t_p - t_o. \end{cases}$$

Die Grössen m_i und σ_i^2 in (35) und (36) hat man sich natürlich durch die ihnen laut (28) entsprechenden Grössen ersetzt vorzustellen.

Wenn wir übrigens in (36) $\sigma_1^2 = \sigma_2^2 = 0$ und $\sigma_3^2 = \sigma_4^2 = \sigma_5^2 = 0$ setzen, so erhalten wir die Streuungen der beiden in (10) angegebenen Verteilungen, nämlich:

$$(37)\quad \begin{cases} \sigma^2(R_{lat_p}) = m_1 m_2 (1-m_2) & \text{für } a \geq t_p - t_o \\ \sigma^2(R_{lat_p}) = m_3 m_4 m_5 (1-m_4 m_5) & \text{für } a < t_p - t_o . \end{cases}$$

Diese trotz exakt gegebener erklärender Grössen verbleibenden Streuungen gehen zurück auf die durch die Verteilungen $s_1(\alpha)$ und γ_1 beschriebenen Ungewissheiten.

Mit Hilfe der Resultate (35) und (36) können wir jetzt auch den Variationskoeffizienten $V(R_{lat_p})$ angeben. Wir setzen hierzu in den Quotienten

$$V^2(R_{lat_p}) = \frac{\sigma^2(R_{lat_p})}{E^2(R_{lat_p})}$$

die Ergebnisse von (35) und (36) ein. Nach einer einfachen Umformung und Einführung der Bezeichnung $V_i = \sigma_i / m_i$ $(i=1,\ldots,5)$[1] für die Variationskoeffizienten der erklärenden Grössen erhalten wir[2]

$$(38)\quad \begin{cases} V^2(R_{lat_p}) = \frac{1}{m_1}(\frac{1}{m_2} - 1 - V_2^2) + V_1^2 + V_2^2 + V_1^2 V_2^2 & \text{für } a \geq t_p - t_o \\ V^2(R_{lat_p}) = \frac{1}{m_3}(\frac{1}{m_4 m_5} - 1 - V_4^2 - V_5^2 - V_4^2 V_5^2) + V_3^2 + V_4^2 + V_5^2 \\ \qquad + V_3^2 V_4^2 + V_3^2 V_5^2 + V_4^2 V_5^2 + V_3^2 V_4^2 V_5^2 & \text{für } a < t_p - t_o . \end{cases}$$

Die für die Existenz der Variationskoeffizienten notwendige Voraussetzung, dass alle m_i von Null verschieden sind, ist im Rahmen unseres Problems gesichert.

---

1) Man vergleiche hierzu wieder die Bezeichnungen (28).

2) Wir geben der Einfachheit halber das Quadrat des Variationskoeffizienten von R_{lat_p} an.

Wenn wir annehmen, dass die Mittelwerte des Ausgangsbestandes $R_{1,a-t_p+t_o,t_o}$ und der Kälbergeburten K_{t_p-a} (das sind die Grössen m_1 und m_3) sehr gross sind, dann können wir durch Bildung des Grenzübergangs "m_1, m_3 gegen Unendlich" die Grösse $V^2(R_{lat_p})$ approximieren durch:

$$(39)\quad \begin{cases} \lim\limits_{m_1 \to \infty} V^2(R_{lat_p}) = V_1^2 + V_2^2 + V_1^2V_2^2 & \text{für } a \geq t_p - t_o \\ \lim\limits_{m_3 \to \infty} V^2(R_{lat_p}) = V_3^2 + V_4^2 + V_5^2 + V_3^2V_4^2 + V_3^2V_5^2 + V_4^2V_5^2 & \\ \qquad + V_3^2V_4^2V_5^2 & \text{für } a < t_p - t_o. \end{cases}$$

Die Relationen (39) bringen als Ergebnis des Gesetzes der grossen Zahl zum Ausdruck, dass bei grossem Ausgangsbestand und hoher Anzahl von Kälbergeburten der Variationskoeffizient von R_{lat_p} nicht mehr von den durch $s_1(\alpha)$ und γ_1 beschriebenen Ungewissheiten mitbestimmt wird, sondern nur noch von den Variationskoeffizienten V_i ($i=1,\ldots,5$) abhängt. Dies stellt eine weitere bereits in §3.4 in Aussicht gestellte Rechtfertigung[1)] einer deterministischen Auffassung des Phänomens dar, wenn ausser den von $s_1(\alpha)$ und γ_1 beschriebenen Ungewissheiten keine weiteren Ungewissheiten auftreten.

Wenn wir zuletzt noch annehmen, dass die Variationskoeffizienten V_i klein gegenüber 1 sind, so lässt sich durch Vernachlässigung der gemischten Produkte $V_i^2V_j^2$ und $V_i^2V_j^2V_k^2$ in (39) die folgende Approximation für $V^2(R_{lat_p})$ angeben:

$$(40)\quad \begin{cases} V^2(R_{lat_p}) \approx V_1^2 + V_2^2 & \text{für } a \geq t_p - t_o \\ V^2(R_{lat_p}) \approx V_3^2 + V_4^2 + V_5^2 & \text{für } a < t_p - t_o. \end{cases}$$

Mit den Ergebnissen dieses Paragraphen sind wir also in der Lage, Mittelwert, Streuung und Variationskoeffizient von R_{lat_p} aus den Mittelwerten, Streuungen und Variationskoeffizienten der erklärenden Grössen zu bestimmen.

---

1) Vgl. Fussnote 1 Seite 33.

§7 DIE OPTIMIERUNG VON INFORMATIONSPROGRAMMEN ZUR VORAUSRECHNUNG VON R_{lat_p}

7.1 Vorbemerkung

Zum Abschluss unserer Untersuchungen wollen wir auf einen wichtigen ökonomischen Aspekt im Zusammenhang mit dem Problem der Vorausrechnung von Rinderprozessen zu sprechen kommen, nämlich auf die Frage, welche statistische Aktivität für die Vorausrechnung die optimale ist.

Auf dieses Problem stossen wir durch die folgende Ueberlegung, welche sich zwangsläufig an die obigen Betrachtungen anschliesst.

Die Streuung von R_{lat_p} wird, laut (36), von den Mittelwerten m_i und den Streuungen σ_i^2 der erklärenden Grössen bestimmt. Während die Mittelwerte von dem Phänomen selbst abhängige, also durch statistische Aktivität nicht beeinflussbare Grössen sind, hängen die Streuungen σ_i^2 von der jeweiligen Informationslage über die erklärenden Grössen ab und können also durch statistische Aktivität gesteuert werden. Je mehr Information man sich beschafft, desto mehr kann man die σ_i^2 einem bestimmten Minimalwert annähern und desto weiter lässt sich dadurch letztlich die Streuung von R_{lat_p} reduzieren. Selbstverständlich ist man bestrebt, diese Streuung so weit wie möglich zu reduzieren, was also die Forderung nach mehr Information zur entsprechenden Reduzierung der einzelnen σ_i^2 beinhaltet. Diese Forderung stösst jedoch auf gewisse Schwierigkeiten in Form der mit der Informationsbeschaffung verbundenen Kosten, welche der Nützlichkeit derselben gegenüberstehen und in Verbindung mit dieser eine bestimmte Grenze definieren, über welche hinaus eine weitere Reduzierung der Streuungen σ_i^2 unrationell ist. Es stellt sich das Problem, diese "Optimalwerte" für die σ_i^2 zu ermitteln, wenn gewisse Vorstellungen über Nutzen und Kosten vorliegen.[1]

1) Diese auf die Streuungen bezogenen Ueberlegungen können analog natürlich auch bezüglich der mit den Streuungen zusammenhängenden Variationskoeffizienten angestellt werden. Wir werden dies nachher auch explizit tun.

Wenn für die Reduzierung der Streuungen σ_i^2 ein gewisses Budget vorgegeben ist, dann besteht das Optimierungsproblem darin, dieses Budget optimal zur Reduzierung der einzelnen σ_i^2 einzusetzen.
Diese beiden Probleme wollen wir im folgenden im Rahmen der allgemeineren Aufgabenstellung der Optimierung sogenannter Informationsprogramme behandeln.

7.2 Das Problem der Optimierung von Informationsprogrammen im allgemeinen[1)]

Wenn Interesse an einem unbekannten Sachverhalt besteht, so versucht man, über denselben Informationen einzuholen, was im Rahmen eines Informationsprogramms geschieht, unter welchem wir ganz allgemein die Gesamtheit aller Grössen und Merkmale verstehen, welche das Phänomen, das die Informationsbeschaffung ermöglicht. eindeutig determinieren. Solche Informationsprogramme haben einerseits einen bestimmten Nutzen oder Wert, andrerseits verbinden sich mit ihnen aber auch bestimmte Kosten, und es besteht deshalb ein besonderes Anliegen, Informationsprogramme in dem Sinn optimal zu gestalten, dass ihr Nutzen und ihre Kosten in einem günstigsten Verhältnis zueinander stehen. Die hierzu vorzunehmende Abwägung zwischen Nutzen und Kosten setzt natürlich in erster Linie die Möglichkeit der Messung derselben voraus. Während die Messung der Kosten prinzipiell als unproblematisch angesehen werden kann, da sich Kosten im allgemeinen direkt in reellen Grössen ausdrücken lassen, ist die Messung des Nutzens weit problematischer, da hierfür meistens erst ein geeignetes Messkonzept aufgestellt werden muss.

Ein zweckmässiger Weg, zu einem Mass für den Nutzen oder Wert eines Informationsprogramms zu gelangen, besteht darin, das betreffende Informationsprogramm im Zusammenhang mit demjenigen Entscheidungsproblem unter Ungewissheit zu sehen, in welchem ihm die Auf-

1) Die folgenden Ausführungen lehnen sich an diesbezügliche Gedankengänge in [6] an.

gabe der Reduzierung einer herrschenden Ungewissheit zukommt. Basierend auf dem ungewissheitsreduzierenden Potential des Informationsprogramms kann in Verbindung mit gewissen Konstituenten des Entscheidungsproblems, insbesondere in Verbindung mit seiner Konsequenzstruktur, ein Mass für den Nutzen des Informationsprogramms konstruiert werden. Auf der Grundlage des betreffenden Entscheidungsmodells kann also die erforderliche Synthese zwischen Nutzen und Kosten vorgenommen werden.

Das beschriebene Problem und den angedeuteten Lösungsweg werden wir jetzt mathematisch formulieren.

7.3 Die allgemeine Lösung des Optimierungsproblems von Informationsprogrammen

Eine unbekannte Grösse θ induziere im Rahmen eines gewissen stochastischen Experimentes e für eine stochastische Veränderliche z[1] eine Wahrscheinlichkeitsverteilung $f(z;\theta|e)$; die durch e bestimmte Funktionalform f sei dabei bekannt. Eine im Rahmen von e entstandene Beobachtung über die Variable z ermöglicht es, über die unbekannte Grösse θ eine statistische Inferenz[2] aufzubauen. Diese Beobachtung - wir bezeichnen sie mit z_o - informiert uns also über θ. Dementsprechend können wir z_o als Information bezeichnen[3] und das Experiment e, welches den (stochastischen) Mechanismus des Zustandekommens der Information z_o beschreibt, als Informationsprogramm.

Wir stellen uns vor, dass wir über eine Menge E solcher Informationsprogramme verfügen, welche uns Informationen über θ liefern können, und dass wir uns in dieser Menge für ein Informationsprogramm e zu entscheiden haben. Woran sollen wir diese Entscheidung orientieren?

---

1) Dieselbe hat man sich noch mit dem Index e versehen vorzustellen, da sie im allgemeinen durch das Experiment e bestimmt wird.

2) Vgl. hierzu §4.2.

3) Die stochastische Veränderliche z heisst dann entsprechend Informationsvariable.

Um das hierfür geeignete Kriterium zu finden, werden wir für die einzelnen Informationsprogramme deren Nutzen und Kosten miteinander in Verbindung bringen.

Was die Kosten betrifft, so gehen wir davon aus, dass über dem Raum E der Informationsprogramme eine Kostenfunktion c(e) gegeben ist, welche jedem e aus E die bei seiner Anwendung zu zahlenden Kosten zuordnet.

Um den Nutzen zu messen, betrachten wir die einzelnen Informationsprogramme im Zusammenhang mit demjenigen Entscheidungsproblem unter Ungewissheit, in welchem Θ den unbekannten Zustand der Natur darstellt und den Informationsprogrammen e aus E die Aufgabe der Reduzierung der über denselben herrschenden Ungewissheit übertragen werden kann.

Wir nehmen an, dass angesichts des unbekannten Zustands der Natur Θ in einem Aktionsraum A eine Entscheidung getroffen werden soll, wobei wir wissen, dass, wenn dieselbe zugunsten der Aktion a aus A ausfällt und sich gleichzeitig der Zustand Θ einstellt, wir bei diesem Zusammentreffen einen Verlust der Höhe $L(a,\Theta)$ in Kauf zu nehmen haben.
Bei rationellem Verhalten ist bei unterschiedlicher Ungewissheit über Θ die Aussicht auf den zu ertragenden Verlust im allgemeinen verschieden . Wegen dieser Disjunktionsleistung wollen wir die Verlustaussicht oder den Erwartungsverlust als Maßstab für den Nutzen des mit der betreffenden Ungewissheit in Zusammenhang stehenden Informationsprogramms betrachten.

Wenn für jedes e eine Verteilung $f(z;\Theta|e)$ gegeben ist und ausserdem für den unbekannten Zustand Θ eine a-priori-Verteilung $p(\Theta)$ bekannt ist, dann stellt der Ausdruck

$$(41) \qquad r^*(z|e) = \inf_{a \in A} \int_{\Theta} L(a,\Theta) \frac{f(z;\Theta|e)p(\Theta)}{\int_{\Theta} f(z;\Theta|e)p(\Theta)d\Theta} d\Theta$$ [1]

zunächst einen Maßstab für den Nutzen[2] einer aus dem Informations-

1) Dieser Ausdruck erinnert an den Begriff "conditional value of the sample information" in [25], Seite 89.
2) Die Grösse $r^*(z|e)$ stellt einen negativen Nutzen dar und wir könnten einen Nutzen $u(z|e)$ definieren durch $u(z|e) = -r^*(z|e)$.

programm e hervorgegangenen Information z dar. Wenn man bezüglich z den Erwartungswert von (41) bildet, so erhält man die Grösse

$$(42) \qquad r(e) = \int_z r^*(z|e) \int_\theta f(z;\theta|e)p(\theta)d\theta \, dz \,.$$ [1]

Diese nur noch von e selbst abhängige Grösse mißt schliesslich den Nutzen des Informationsprogramms e; je kleiner r(e) ist, desto grösser ist der Nutzen von e und umgekehrt.

Um das gesuchte Kriterium für unsere Entscheidung in E zu erhalten, bringen wir den Nutzen und die Kosten von e in der Weise in Verbindung miteinander, dass wir zu der Grösse r(e) die Kosten c(e) addieren. Die Summe

$$(43) \qquad F(e) = r(e) + c(e)$$

stellt dann einen Maßstab für den Effektiv-Nutzen [2] von e dar und liefert das gesuchte Kriterium zur Auffindung des optimalen Informationsprogramms e_o in E.

Wir bezeichnen dasjenige Informationsprogramm e_o in E als optimal, für welches die Grösse F(e) ihren kleinsten Wert annimmt.

Durch Einsetzen dieses optimalen e_o in die Kostenfunktion c(e) erhalten wir schliesslich die in der betreffenden Situation aufzubringenden Optimalkosten $c(e_o)$.

Die Bestimmung des optimalen e_o erfordert natürlich je nach Beschaffenheit des Raumes E unterschiedliche Verfahren.
Für analytische Verfahren zur Minimierung von F(e) ist es zweckmässig, den Raum E und seine Elemente e in dem Sinn zu strukturieren, dass festgestellt wird, welche nutzen- oder kostenrelevante Faktoren die in E vertretenen Informationsprogramme eindeutig beschreiben. Wenn die Gesamtheit all dieser Faktoren in d Grössen β_i (i=1,...,d) besteht, dann sagen wir die Informationsprogramme e aus E sind von der Struktur $(\beta_1,\ldots,\beta_d)$. Die Grössen β_i selbst kann man als Bestimmungsparameter, Strukturparameter oder in Anbetracht der Tatsache, dass sie Gegenstand einer Optimierungsoperation sind, als Operationsparameter von e ansprechen.

---

1) Diese Grösse erinnert an den Begriff "expected value of the sample information" in [25], Seite 90.
2) Diese Grösse erinnert an den Begriff "expected net gain of sampling of an experiment" in [25], Seite 91.

Wenn alle Strukturparameter $\beta_1,\ldots,\beta_d$ der Informationsprogramme e aus E als reelle Grössen aufgefasst werden können, dann kann E mit einem Teilraum des Euklidischen Raumes von d Dimensionen identifiziert werden. Die Punkte dieses Teilraumes stellen dann je ein Informationsprogramm $e=(\beta_1,\ldots,\beta_d)$ dar. F(e) ist dann eine reelle Funktion $F(\beta_1,\ldots,\beta_d)$ über diesem Teilraum und bei Gegebenheit gewisser Voraussetzungen bezüglich F und E[1] kann man das optimale $e_o = (\beta_{1o},\ldots,\beta_{do})$ als Lösung des Gleichungssystems

$$\frac{\partial F(\beta_1,\ldots,\beta_d)}{\partial \beta_i} = 0 \qquad i = 1,\ldots,d \tag{44}$$

finden.

Beim obigen Ansatz zum Optimierungsproblem von Informationsprogrammen werden, wenn c(e) eine nichtkonstante Kostenfunktion ist und der Raum E somit mehrere Kostenalternativen zulässt, gleichzeitig die Optimalkosten und deren optimale Verwendung bestimmt.
In der Praxis kommt es aber häufig vor, dass von vornherein die Kosten dadurch festgelegt sind, dass ein gewisses Budget C_o für die Informationsgewinnung bereitgestellt ist. Die damit gegebene Budget-Bedingung $c(e) = C_o$ definiert in E einen Teilraum $E(C_o)$, welchem alle e aus E mit der Eigenschaft $c(e) = C_o$ angehören, und es besteht nur noch das Problem, dasjenige Informationsprogramm in $E(C_o)$ zu finden, welches die Grösse r(e) minimiert[2] und damit die optimale Verwendung des Budgets C_o garantiert.

---

1) Wegen des Problems der Bestimmung von Extrema von Funktionen mehrerer Veränderlicher und der entsprechenden Voraussetzungen vgl.[23], Band 2, Seite 274 ff.

2) Dies ist mathematisch gesehen gleichbedeutend damit, F(e) in E unter Beachtung der Nebenbedingung $c(e) = C_o$ zu minimieren.

Den Begriff des Informationsprogramms hatten wir oben in sehr abstrakter Form eingeführt. Das folgende einfache Beispiel möge dazu dienen, den Inhalt dieses Begriffs zu erläutern und insbesondere zu zeigen, welche Grössen z.B. als Operationsparameter eines Informationsprogramms fungieren können.

Der Mittelwert Θ einer normal-verteilten Grösse x mit bekannter Streuung σ_o^2 sei unbekannt und soll geschätzt werden.
Es stehe uns offen, zu diesem Zweck n (n=1,2,3,...) unabhängige Zufallsbeobachtungen über die Grösse x vorzunehmen, wobei allerdings jede Beobachtung x_i (i=1,...,n) einen Zufallsfehler ε_i enthält, welcher sie vom wirklichen Wert x_i^* um ε_i abweichen lässt: $x_i - x_i^* = \varepsilon_i$ (i=1,...,n). Diese Zufallsfehler mögen den Mittelwert $E(\varepsilon_i) = 0$ und die Streuung $\sigma^2(\varepsilon_i) = \sigma_\varepsilon^2$ haben. Wir nehmen an, dass wir, ebenso wie wir über die Anzahl n der Beobachtungen entscheiden können, auch über die Grösse σ_ε^2 entscheiden können, indem wir m verschiedene Beobachtungsmethoden oder Instrumente zur Auswahl haben und wissen, dass die Methode μ (μ=1,...,m) zur Streuung $\sigma_{\varepsilon\mu}^2$ des Zufallsfehlers der Beobachtung führt.
Die Grösse n, welche also die Quantität an Informationen angibt, und die Grösse σ_ε^2, welche die Qualität derselben bestimmt, determinieren also eindeutig unsere Informationslage über Θ. Sie sind die Bestimmungsparameter der möglichen Informationsprogramme zur Schätzung von Θ - durch Angabe von n und σ_ε^2 ist ein Informationsprogramm über Θ gegeben. Der Raum der möglichen Informationsprogramme ist gleich der Gesamtheit der Vektoren $(n, \sigma_\varepsilon^2)$, für welche n eine natürliche Zahl und σ_ε^2 ein Element der Menge $(\sigma_{\varepsilon 1}^2, \ldots, \sigma_{\varepsilon m}^2)$ ist.
Ein Informationsprogramm $(n, \sigma_\varepsilon^2)$ induziert dann für die als Informationsvariable betrachtete Grösse $\frac{1}{n}\sum x_i$ eine Normalverteilung mit Mittelwert Θ und Streuung $(\sigma_o^2 + \sigma_\varepsilon^2)/n$.

Zur Ermittlung des optimalen Informationsprogramms bedarf es noch der Kenntnis einer Kostenfunktion $c(n, \sigma_\varepsilon^2)$ sowie der Präzisierung des Entscheidungsproblems, in dessen Rahmen die Information über Θ von Bedeutung ist. Wie bei solchen Angaben weiter zu verfahren ist, wird in den folgenden Ausführungen ersichtlich.

7.4 Die Optimierung von Informationsprogrammen zur Vorausrechnung von R_{lat_p}

Das soeben dargestellte Konzept wollen wir jetzt auf unser Problem der Vorausrechnung von R_{lat_p} anwenden.

Die uns interessierende unbekannte Grösse, über welche wir uns informieren wollen, ist R_{lat_p}. Informationen über R_{lat_p} können allerdings nur indirekt auf die Art zustande kommen, dass Informationen über die erklärenden Grössen von R_{lat_p} eingeholt werden und dieselben an Hand des Modells (9,10,11,12)[1] in Information über R_{lat_p} transformiert werden. Informationsprogramme für R_{lat_p} stellen also im Grunde Gesamtheiten von Einzelinformationsprogrammen über die erklärenden Grössen dar.

Wir bezeichnen Informationsprogramme

$$(45)\qquad \begin{cases} \text{über } R_{1,a-t_p+t_o,t_o} & \text{mit } e_1 \\ \text{über } p_{lat_p t_o} & \text{mit } e_2 \\ \text{über } K_{t_p-a} & \text{mit } e_3 \\ \text{über } q_{1a} & \text{mit } e_4 \\ \text{über } \gamma_1 & \text{mit } e_5. \end{cases}$$

Informationsprogramme über R_{lat_p} sind dann gegeben durch

$$(46)\qquad \begin{cases} \text{die Gesamtheit } (e_1,e_2) & \text{für } a \geq t_p-t_o \\ \text{die Gesamtheit } (e_3,e_4,e_5) & \text{für } a<t_p-t_o. \end{cases}$$

Wir nehmen an, es seien je ein Raum E_{12} bzw. E_{345} von Informationsprogrammen (e_1,e_2) bzw. (e_3,e_4,e_5) gegeben. Die durch solche Informationsprogramme vermittelten Informationen über R_{lat_p} bezeichnen wir mit (z_1,z_2) bzw. (z_3,z_4,z_5).

Ein Informationsprogramm e und aus ihm resultierende Information z[1] induzieren eine Verteilung $\omega(R_{lat_p}; z|e)$. Die Kostenfunktion c(e) über dem jeweiligen Raum der Informationsprogramme e sei bekannt.

---

1) Wir schreiben abkürzend e anstelle von (e_1,e_2) bzw. (e_3,e_4,e_5) und z anstelle von (z_1,z_2) bzw. (z_3,z_4,z_5).

Zur Messung des Nutzens der Informationsprogramme e stellen wir R_{lat_p} wieder in den in § 5 geschilderten entscheidungstheoretischen Zusammenhang und betrachten R_{lat_p} als unbekannten Zustand und die möglichen Prognosen $\hat{R}_{lat_p}$ als die zur Auswahl stehenden Aktionen. Bezüglich der Verlustfunktion dieses Entscheidungsproblems greifen wir auf die in § 5.5 getroffene Vereinbarung zurück und betrachten nacheinander die beiden dort festgelegten Ansätze.

A) Die Verlustfunktion $L_1(\hat{R},R)^{1)} = k(\hat{R}-R)^2$:

Mit dieser Verlustfunktion erhalten wir gemäss (41) als Maßstab des Nutzens von z die Grösse

$$(47) \qquad r^*(z|e) = \inf_{\hat{R}} \int_R k(\hat{R}-R)^2 \omega(R;z|e)dR^{2)}.$$

Laut § 5.3.a ist (47) gleich $k\cdot\sigma^2(R)$. Die Streuung $\sigma^2(R)$ wurde in (36) als Funktion der Streuungen σ_i^2 und der Mittelwerte m_i der erklärenden Grössen ausgedrückt. Die σ_i^2 und m_i hängen im allgemeinen von der Information z_i ab, so dass wir, um von $r^*(z|e)$ zum Nutzenmaß $r(e)$ zu gelangen, für $\sigma^2(R)$ noch den Erwartungswert bezüglich z zu bilden hätten.

Der Einfachheit halber nehmen wir jedoch die m_i als vorherbestimmte Grössen und die σ_i^2 als von z unabhängige Grössen an, so dass $k\cdot\sigma^2(R)$ gleichzeitig auch das Nutzenmaß $r(e)$ für e darstellt.[3)]

---

1) Zur Abkürzung werden die Indizes lat_p wieder vernachlässigt.

2) Dies ist das Bayessche Risiko in der in § 5.2 erwähnten Eigenschaft als Maß des Wertes der über R vorliegenden Information.

3) Wenn diese Annahme nicht erfüllt ist und die Bildung des Erwartungswertes bezüglich z erforderlich ist, bleibt der resultierende Ansatz formell derselbe. Man braucht sich dann lediglich die m_i und σ_i^2 durch ihre Erwartungswerte bezüglich z ersetzt vorzustellen und nur zu verlangen, dass $E_z(m_i)$ von e unabhängig ist. Um diese Ersetzung vornehmen zu können, ist nur zu fordern, dass die Informationen z_i bezüglich der einzelnen Parameter stochastisch unabhängig anfallen, was in der Praxis zumindest näherungsweise gewährleistet sein dürfte. Diesen wichtigen Hinweis verdanke ich Herrn Prof.Dr.H.Schneeweiß.

Wegen (36) bemißt sich der Nutzen von e also an der Grösse

$$(48)\quad r(e)=\begin{cases} k[m_1m_2-m_1^2m_2^2+(\sigma_1^2+m_1^2-m_1)(\sigma_2^2+m_2^2)] & \text{für } a\geq t_p-t_o \\ k[m_3m_4m_5-m_3^2m_4^2m_5^2+(\sigma_3^2+m_3^2-m_3)(\sigma_4^2+m_4^2)(\sigma_5^2+m_5^2)] & \text{für } a<t_p-t_o. \end{cases}$$

Die Mittelwerte m_i sind als Operationsparameter von e irrelevant.[1] Die Streuungen σ_i^2 dagegen können als Funktionen nicht näher spezifizierter elementarer Operationsparameter von e als "aggregierte" Operationsparameter von e aufgefasst werden. Unter Verzicht auf die Kenntnis der elementaren Operationsparameter von e, betrachten wir die σ_i^2 als charakteristische Grössen von e und identifizieren ein Informationsprogramm e schliesslich mit der Gesamtheit der durch seine elementaren Bestimmungsparameter gestifteten Streuungen σ_i^2.

Wir definieren also

$$(49)\qquad e=\begin{cases} (e_1,e_2)=(\sigma_1^2,\sigma_2^2) & \text{für } a\geq t_p-t_o \\ (e_3,e_4,e_5)=(\sigma_3^2,\sigma_4^2,\sigma_5^2) & \text{für } a<t_p-t_o. \end{cases}$$

Wir vereinbaren damit also, dass wir ein Informationsprogramm e zur Vorausrechnung von R_{lat_p} als eindeutig beschrieben ansehen, wenn wir die Streuungen σ_1^2 und σ_2^2 bzw. σ_3^2, σ_4^2 und σ_5^2 kennen.

Was die Kosten c(e) betrifft,[2] so hängen dieselben zwar auch von gewissen elementaren Parametern von e ab, aber wir unterstellen, dass sie sich auch als Funktion der soeben als Operationsparameter erklärten Streuungen σ_i^2 ausdrücken lassen, so dass wir schreiben können

$$(50)\qquad c(e)=\begin{cases} c(e_1,e_2)=c(\sigma_1^2,\sigma_2^2) & \text{für } a\geq t_p-t_o \\ c(e_3,e_4,e_5)=c(\sigma_3^2,\sigma_4^2,\sigma_5^2) & \text{für } a<t_p-t_o. \end{cases}$$

---

1) Da wir sie erstens soeben als vorherbestimmte Grössen deklariert haben, und da sie zweitens ohnehin dem Phänomen inhärente und damit nicht steuerbare Grössen sind.

2) Wegen allgemeiner Untersuchungen über Kostenfunktionen im Bereich der Agrarstatistik vgl. z.B. [13].

Die Kostenfunktion wird normalerweise eine für abnehmende σ_i^2 monoton steigende Funktion darstellen.

Gewisse a-priori-Beschränkungen bezüglich des Raumes E der möglichen Informationsprogramme kann man übrigens durch geeignete Festlegung des Definitionsbereiches von c(e) berücksichtigen. Wenn die Streuungen σ_i^2 beispielsweise bestimmte Werte σ_i^{*2} weder unterschreiten noch erreichen können, so beschränkt man den Definitionsbereich von c(e) auf $E_{12}=\{\sigma_i^2 > \sigma_i^{*2}; i=1,2\}$ bzw. $E_{345}=\{\sigma_i^2 > \sigma_i^{*2}; i=3,4,5\}$. Dies geschieht z.B. im folgenden Ansatz einer Kostenfunktion:

$$(51)\quad \begin{cases} c(\sigma_1^2,\sigma_2^2) = \dfrac{c_1}{(\sigma_1^2-\sigma_1^{*2})^{h_1}} + \dfrac{c_2}{(\sigma_2^2-\sigma_2^{*2})^{h_2}}, & c_i,h_i>0,\ \sigma_i^2>\sigma_i^{*2}, \\ & (i=1,2), \\ c(\sigma_3^2,\sigma_4^2,\sigma_5^2) = \dfrac{c_3}{(\sigma_3^2-\sigma_3^{*2})^{h_3}} + \dfrac{c_4}{(\sigma_4^2-\sigma_4^{*2})^{h_4}} + \dfrac{c_5}{(\sigma_5^2-\sigma_5^{*2})^{h_5}}, & \\ & c_i,h_i>0,\ \sigma_i^2>\sigma_i^{*2},(i=3,4,5). \end{cases}$$

Zur Beurteilung des Effektiv-Nutzens von e orientieren wir uns gemäss dem Ausdruck (43) an den folgenden Grössen:

$$(52)\quad \begin{cases} k[m_1m_2-m_1^2m_2^2+(\sigma_1^2+m_1^2-m_1)(\sigma_2^2+m_2^2)] + \dfrac{c_1}{(\sigma_1^2-\sigma_1^{*2})^{h_1}} + \dfrac{c_2}{(\sigma_2^2-\sigma_2^{*2})^{h_2}} & \\ & \text{für } a\geq t_p-t_o, \\ k[m_3m_4m_5-m_3^2m_4^2m_5^2+(\sigma_3^2+m_3^2-m_3)(\sigma_4^2+m_4^2)(\sigma_5^2+m_5^2)] + \dfrac{c_3}{(\sigma_3^2-\sigma_3^{*2})^{h_3}} & \\ + \dfrac{c_4}{(\sigma_4^2-\sigma_4^{*2})^{h_4}} + \dfrac{c_5}{(\sigma_5^2-\sigma_5^{*2})^{h_5}} & \text{für } a<t_p-t_o. \end{cases}$$

Der Raum von Informationsprogrammen, über welchem (52) zu betrachten ist, ist gegeben durch $E_{12}=\{\sigma_i^2>\sigma_i^{*2}; i=1,2\}$ bzw. $E_{345}=\{\sigma_i^2>\sigma_i^{*2}; i=3,4,5\}$. Die optimalen Informationsprogramme (e_{1o},e_{2o})[1] bzw. (e_{3o},e_{4o},e_{5o}) zur Vorausrechnung von R_{lat_p} sind durch diejenigen Grössen $(\sigma_{1o}^2,\sigma_{2o}^2)$ bzw. $(\sigma_{3o}^2,\sigma_{4o}^2,\sigma_{5o}^2)$ gegeben, welche (52) minimieren.

---

1) Der Index o soll die Optimalität kennzeichnen.

Man kann dieselben durch Auflösung des Gleichungssystems

$$(53a') \quad \begin{cases} k(\sigma_2^2 + m_2^2) - \dfrac{h_1 c_1}{(\sigma_1^2 - \sigma_1^{*2})^{h_1+1}} = 0 \\ k(\sigma_1^2+m_1^2-m_1) - \dfrac{h_2 c_2}{(\sigma_2^2 - \sigma_2^{*2})^{h_2+1}} = 0 \end{cases}$$

nach (σ_1^2, σ_2^2), bzw. durch Auflösung des Gleichungssystems

$$(53b) \quad \begin{cases} k(\sigma_4^2+m_4^2)(\sigma_5^2+m_5^2) - \dfrac{h_3 c_3}{(\sigma_3^2 - \sigma_3^{*2})^{h_3+1}} = 0 \\ k(\sigma_3^2+m_3^2-m_3)(\sigma_5^2+m_5^2) - \dfrac{h_4 c_4}{(\sigma_4^2 - \sigma_4^{*2})^{h_4+1}} = 0 \\ k(\sigma_3^2+m_3^2-m_3)(\sigma_4^2+m_4^2) - \dfrac{h_5 c_5}{(\sigma_5^2 - \sigma_5^{*2})^{h_5+1}} = 0 \end{cases}$$

nach $(\sigma_3^2, \sigma_4^2, \sigma_5^2)$ finden.

Die Lösung dieser Gleichungssysteme kann nicht in geschlossener Form angegeben werden, da bei der Auflösung Gleichungen höheren als zweiten Grades auftreten. Die Lösungen sind mit numerischen Verfahren zu ermitteln.

Durch Einsetzen der gefundenen Lösungen σ_{io}^2 in die Gleichungen (36) erhält man die Optimalstreuung für R_{lat_p} und durch Einsetzen in die Kostenfunktion (51) die entsprechenden Optimalkosten.

Für den Fall, dass bezüglich der Kostenfunktion (51) die Bedingungen

$$(54) \quad \begin{cases} c(\sigma_1^2, \sigma_2^2) = C_o \\ c(\sigma_3^2, \sigma_4^2, \sigma_5^2) = C_{oo} \end{cases}$$

vorliegen, hat man zur Auffindung des optimalen Informationsprogramms die Ausdrücke (48) über den entsprechenden durch (54) definierten Hyperräumen zu minimieren[1]. Die ebenfalls numerisch zu ermittelnden Lösungen[2] $\sigma^2_{io}(C_o)$ bzw. $\sigma^2_{io}(C_{oo})$ beschreiben dasjenige Informationsprogramm, welches die optimale Verwendung des durch (54) festgelegten Budgets C_o bzw. C_{oo} zur Vorausrechnung von R_{lat_p} garantiert.

Schliesslich wollen wir noch einen wichtigen Teilraum im Raum der Informationsprogramme definieren:

> Die Gesamtheit aller Informationsprogramme, welche für irgend ein k (k>o) bzw. irgend ein (C_o, C_{oo}) optimal sind, bezeichnen wir als den Raum der zulässigen Informationsprogramme. Alle nicht diesem Raum angehörenden Informationsprogramme sind insofern unrationell (und deshalb unzulässig), als zu ihnen bei gleichen Kosten ein Programm mit grösserem Nutzen, bzw. bei gleichem Nutzen ein Programm mit niedrigeren Kosten gefunden werden kann.

In unserem Fall ist der Raum der zulässigen Informationsprogramme beschrieben durch die Gleichungssysteme (53a) bzw. (53b), wenn man dort die Grösse k als variablen Parameter betrachtet. Die durch diese Gleichungssysteme definierten Hyperräume stellen jeweils eine Kurve im 2- bzw. 3-dimensionalen Raum dar, denn die von k abhängigen Lösungen $(\sigma^2_{1o}(k), \sigma^2_{2o}(k))$ bzw. $(\sigma^2_{3o}(k), \sigma^2_{4o}(k), \sigma^2_{5o}(k))$ von (53a) bzw. (53b) liefern bei Betrachtung von k als variablem Parameter je eine Parameterdarstellung einer Kurve.[3]

Diese letzten Ueberlegungen lassen sich wegen der Nichtdarstellbarkeit der Lösungen σ^2_{io} leider analytisch nicht näher ausführen, wir werden dies aber bei Verwendung der zweiten Verlustfunktion tun können.

---

1) Dies ist das allgemeine Problem der Minimierung von Funktionen mehrerer Veränderlicher bei gegebenen Nebenbedingungen. Man vergleiche hierzu [23], Band 2, Seite 281 ff.

2) Diese Lösungen hängen von C_o bzw. C_{oo} ab, sind aber von k unabhängig, denn k stellt in den zu minimierenden Ausdrücken (48) nur einen konstanten Faktor dar, welcher keinen Einfluss auf die Lage des Minimums hat.

3) Man vergleiche hierzu [23], Band 2, Seite 347 ff.

B) Die Verlustfunktion $L_2(\hat{R},R) = k(\hat{R}-R)^2/\hat{R}^2$:

Bei Zugrundelegung dieser Verlustfunktion ist analog zu (47) der Maßstab für den Nutzen von z gegeben durch die Grösse

$$(55) \qquad r^*(z|e) = \inf_{\hat{R}} \int_R k\left(\frac{\hat{R}-R}{\hat{R}}\right)^2 \omega(R;z|e)dR$$

Laut §5.3.f ist (55) gegeben durch die Grösse $kV^2(R)/(V^2(R)+1)$, welche wir, genau wie wir es unter A) getan haben, gleichzeitig als Maßstab des Nutzens von e ansehen.

Zum Zwecke der Vereinfachung unserer Betrachtungen nehmen wir an, dass die Räume E_{12} bzw. E_{345} nur Informationsprogramme enthalten, welche bereits so kleine Variationskoeffizienten V(R) liefern, dass wir die Grösse $kV^2(R)/(V^2(R)+1)$ durch $kV^2(R)$ ausreichend gut approximieren können.
Da wir ferner m_1 und m_3 als so gross ansehen können, dass $V^2(R)$ durch (39) hinreichend gut approximiert wird, und wir ausserdem, da mit V(R) auch die elementaren V_i klein sind, auch die gemischten Produkte in (39) vernachlässigen und (40) als Approximation von $V^2(R)$ verwenden können, besitzen wir schliesslich mit

$$(56) \qquad k(V_1^2 + V_2^2) \quad \text{bzw.} \quad k(V_3^2 + V_4^2 + V_5^2)$$

ein geeignetes Mass für den Nutzen von e.

Analog zu oben fassen wir die V_i als Operationsparameter der Informationsprogramme e auf und identifizieren ein Informationsprogramm e mit einem Vektor (V_1,V_2) bzw. (V_3,V_4,V_5).

Wenn wir die Kostenfunktion c(e) ebenfalls als Funktion der V_i spezifizieren und beispielsweise den einfachen Ansatz

$$(57) \qquad \begin{cases} c(V_1,V_2) = \dfrac{c_1}{V_1^{h_1}} + \dfrac{c_2}{V_2^{h_2}} \;, & c_i,h_i > 0,\ V_i > 0,\ (i=1,2), \\ c(V_3,V_4,V_5) = \dfrac{c_3}{V_3^{h_3}} + \dfrac{c_4}{V_4^{h_4}} + \dfrac{c_5}{V_5^{h_5}} \;, & c_i,h_i > 0,\ V_i > 0, \\ & \qquad (i=3,4,5). \end{cases}$$

wählen, dann bemisst sich der Effektiv-Nutzen von e schliesslich an der Grösse

$$(58)\qquad \begin{cases} k(V_1^2 + V_2^2) + \dfrac{c_1}{V_1^{h_1}} + \dfrac{c_2}{V_2^{h_2}} & \text{für } a \geq t_p - t_o \\ k(V_3^2 + V_4^2 + V_5^2) + \dfrac{c_3}{V_3^{h_3}} + \dfrac{c_4}{V_4^{h_4}} + \dfrac{c_5}{V_5^{h_5}} & \text{für } a < t_p - t_o, \end{cases}$$

welche definiert ist über $E_{12}=\{V_i > 0; i=1,2\}$ bzw. $E_{345}=\{V_i > 0; i=3,4,5\}$. Die Ausdrücke in (58) nehmen, wie man durch Bildung der ersten Ableitung nach V_i und Nullsetzen derselben leicht nachprüfen kann, ihre Minima für

$$(59)\qquad V_{io} = \left[\frac{h_i c_i}{2k}\right]^{\frac{1}{h_i+2}}, \quad (i=1,2) \text{ bzw. } (i=3,4,5),$$

an. Das optimale Informationsprogramm ist also durch (V_{1o}, V_{2o}) bzw. (V_{3o}, V_{4o}, V_{5o}) gegeben.[1)]

Wenn wir in (59) die Grösse k als variablen Parameter auffassen ($o<k<\infty$), so liefert (59) die Parameterdarstellung einer Kurve im positiven Bereich der (V_1, V_2)-Ebene bzw. einer Kurve im positiven Bereich des (V_3, V_4, V_5)-Raumes. Diese Kurven repräsentieren den jeweiligen Raum der zulässigen Informationsprogramme.

Bei Vorgabe einer Kosten-Bedingung

$$(60)\qquad \begin{cases} \dfrac{c_1}{V_1^{h_1}} + \dfrac{c_2}{V_2^{h_2}} = C_o, & c_i, h_i > 0,\ V_i > 0 \quad (i=1,2), \\ \dfrac{c_3}{V_3^{h_3}} + \dfrac{c_4}{V_4^{h_4}} + \dfrac{c_5}{V_5^{h_5}} = C_{oo}, & c_i, h_i > 0,\ V_i > 0 \quad (i=3,4,5), \end{cases}$$

ergibt sich das optimale Informationsprogramm als Schnittpunkt

---

1) Durch Einsetzen der durch (59) gegebenen Grössen V_{io} in (40) erhält man den (approximativ) optimalen Variationskoeffizienten für R_{lat_p} und durch Einsetzen der V_{io} in (57) die damit verbundenen Optimalkosten.

zwischen dem jeweiligen durch (60) definierten Hyperraum mit dem entsprechenden oben beschriebenen Raum der zulässigen Informationsprogramme.[1] Wenn wir $h_i = h$ setzen (für alle i), dann können wir diesen Schnittpunkt analytisch ermitteln und erhalten

$$V_{io} = c_i^{\frac{1}{h+2}} \left[\frac{c_1^{\frac{2}{h+2}} + c_2^{\frac{2}{h+2}}}{C_o} \right]^{\frac{1}{h}} \qquad \text{für } i = 1,2$$

(61)

$$V_{io} = c_i^{\frac{1}{h+2}} \left[\frac{c_3^{\frac{2}{h+2}} + c_4^{\frac{2}{h+2}} + c_5^{\frac{2}{h+2}}}{C_{oo}} \right]^{\frac{1}{h}} \qquad \text{für } i = 3,4,5.$$

Für den Fall, dass alle h_i gleich sind und die anderen unseren Untersuchungen zugrunde gelegten Annahmen Gültigkeit haben, beschreibt (61) dasjenige Informationsprogramm, welches das auf Grund der Relationen (60) festgelegte Budget C_o bzw. C_{oo} in optimaler Weise für die Informationsbeschaffung zur Vorausrechnung von R_{lat_p} einsetzt.

---

1) Die Schnittmengen dieser Räume sind Punkte, denn es handelt sich um den Schnitt zweier Kurven, welche offensichtlich nur einen Punkt gemeinsam haben, bzw. um den Schnitt einer Kurve mit einer Fläche, welche sich ebenfalls nur in einem Punkt treffen.

§ 8 ZUSAMMENFASSUNG

Abschliessend wollen wir uns nochmals die einzelnen Schritte unserer Untersuchungen und deren wesentliche Ergebnisse in Erinnerung bringen.

Nach den allgemeinen Einführungen des § 1 haben wir im § 2 im Anschluss an die einleitenden phänomenologischen Betrachtungen und Definitionen unserer Grundbegriffe, insbesondere des Rinderbestandes, der Bestandsmerkmale Laufbahn und Alter sowie des Rinderprozesses, zunächst geklärt, welche Faktoren für die Fortschreibung eines Rinderprozesses erforderlich sind : nämlich der Ausgangsbestand sowie die Neuzugänge und Schlachtungen in den einzelnen Laufbahnen und Altersgruppen. Auf diesen Bestimmungsfaktoren bauten wir die Fortschreibungsformel (1) auf, welche der Fortschreibung eines Rinderprozesses dient, wenn Angaben über alle Bestimmungsfaktoren vorliegen.

Im § 3 entwickelten wir in der Formel (3) ein deterministisches Modell zur Vorausrechnung von Rinderprozessen, wobei in diesem Modell gegenüber der Fortschreibungsformel (1) die Elemente Schlachtungen und Neuzugänge gegen die Begriffe Schlachtaltersverteilung, Kälbergeburten und Laufbahnverteilung ausgetauscht wurden.

Aus dem zentralen Modell (3) konnten noch einige interessante Ableitungen vorgenommen werden.

Eine Diskussion der Voraussetzungen des Modells sowie der Verfügbarkeit bzw. Bestimmbarkeit seiner Daten liess keine besonderen Bedenken aufkommen, hingegen belehrte eine kritische Betrachtung bezüglich der Leistung des Modells, dass dasselbe wegen der zugrundeliegenden deterministischen Auffassung nur wenigen Anforderungen genügen kann, und dass für gewisse Ansprüche die Konzeption eines stochastischen Modells vonnöten ist.

Dieses stochastische Modell entstand in § 4 und zwar in der Weise, dass wir in das Konstruktionsprinzip des Modells (3) potentielle stochastische Elemente des Phänomens mit einbezogen. Das stochastische Modell, welches in der Formel (9) und deren Zusatz (11) sowie den Formeln (10) und (12) besteht, erklärt den zukünftigen(stochastischen)Rinderbestand R_{lat_p} in Abhängigkeit der (stochastischen)Erklärungsgrössen Ausgangsbestand, Schlachtaltersverteilung, Kälbergeburten und Laufbahnverteilung, sowie die Wahrscheinlichkeitsverteilung $\omega(R_{lat_p})$ in Abhängigkeit von den Wahrscheinlichkeitsverteilungen φ_i dieser erklärenden Grössen. Wegen der Kompliziertheit dieses letzten Zusammenhangs, welche uns die Relation (12) offenbarte, liessen wir vorerst von weiteren Untersuchungen am Objekt der Verteilung $\omega(R_{lat_p})$ ab und behandelten im zweiten Teil des § 4 lediglich noch das Problem der Konstruktion von Wahrscheinlichkeitsverteilungen φ_i für die erklärenden Grössen des Modells mit Hilfe der Bayesschen Inferenzmethode.

Zu wichtigen Erkenntnissen führte uns § 5, wo wir von einer formell als gegeben angenommenen Verteilung $\omega(R_{lat_p})$ ausgingen und zur Beantwortung der Frage nach der optimalen Prognose $\hat{R}_{lat_p}$ für die unbekannte Grösse R_{lat_p} eine entscheidungstheoretische Auffassung zugrunde legten. Im Zusammenhang mit der Untersuchung dieses Problems ergab sich gleichzeitig eine elegante Lösung für die Schwierigkeit der Ermittlung der Verteilung $\omega(R_{lat_p})$ aus den Elementarverteilungen φ_i. Denn aus der entscheidungstheoretischen Sicht kristallisierten sich für das Problem der Prognose von R_{lat_p} nur zwei Grössen als allein relevant heraus : die durch (22) gegebene Bayessche Prognose als optimale Prognose für R_{lat_p} und das damit verbundene, ebenfalls durch (22) gegebene Bayessche Risiko als Maßstab der Tragweite der über R_{lat_p} herrschenden Ungewißheit.

Dies bedeutet, dass aus der entscheidungstheoretischen Perspektive nur diejenigen Parameter der Verteilung $\omega(R_{lat_p})$ von Bedeutung sind, mit welchen sich diese beiden Grössen angeben lassen - diese Parameter enthalten also alle in der betreffenden Situation relevante Information über R_{lat_p}. Die Untersuchungen in § 5.3 zeigten, von welchen Parametern der Verteilung $\omega(R_{lat_p})$ bei verschiedenen Verlustfunktionen die Bayessche Prognose und (in einigen Fällen) das Bayessche Risiko bestimmt wird. Von besonderem Interesse waren dabei die Ergebnisse von § 5.3.a und § 5.3.f , dass nämlich bei Zugrundelegung der Verlustfunktionen $L(\hat{R},R)=k(\hat{R}-R)^2$ bzw. $L(\hat{R},R) = k\ (\hat{R}-R)^2/\ \hat{R}^2$ der Erwartungswert und die Streuung bzw. der Variationskoeffizient von R_{lat_p} die einzigen Parameter von $\omega(R_{lat_p})$ sind, welche man zur Angabe der Bayesschen Prognose und des Bayesschen Risikos kennen muss. Wegen der sich mit ihnen verbindenden Vereinfachung der Problematik wurden diese beiden Verlustfunktionen für unsere weiteren Untersuchungen festgelegt.

Im § 6 stellten wir den Erwartungswert und die Streuung sowie den Variationskoeffizienten von R_{lat_p} in Abhängigkeit der Erwartungswerte, Streuungen und Variationskoeffizienten der erklärenden Grössen von R_{lat_p} dar. Die Ergebnisse (35) - (40) des § 6 stellen den für unsere Belange erheblichen Inhalt der komplizierten Relation (12) dar.

Im § 7 beschäftigte uns der ökonomische Aspekt der Vorausrechnung von R_{lat_p}, nämlich das Problem der Rationalisierung des mit ihr verbundenen statistischen Aufwandes. Bei diesem Problem ging es darum, die Aktionsgrössen der statistischen Aktivität, als welche wir letztlich die Streuungen bzw. Variationskoeffizienten der erklärenden Grössen von R_{lat_p} ansahen, in dem Sinn zu optimieren, dass ihr Nutzeneffekt und

ihre Kosten im günstigsten Verhältnis zueinander stehen.

Wir fassten dieses Problem als ein Problem der Optimierung von Informationsprogrammen auf, für welches wir zunächst das auf einem entscheidungstheoretischen Ansatz beruhende theoretische Konzept samt Lösungsgang darstellten, um dasselbe schliesslich auf unser spezielles Problem anzuwenden. Dabei legten wir hinsichtlich des als Basis dienenden Entscheidungsmodells nacheinander die beiden oben vereinbarten Verlustfunktionen zugrunde.

Bei Anwendung der Verlustfunktion $L(\hat{R},R)=k(\hat{R}-R)^2$ konnte der Begriff des Informationsprogramms mit der Gesamtheit der Streuungen der erklärenden Grössen von R_{lat_p} identifiziert und für das optimale Informationsprogramm schliesslich in Form von (53a) bzw. (53b) ein allerdings nicht analytisch lösbares Gleichungssystem angegeben werden.

Bei Zugrundelegung der Verlustfunktion $L(\hat{R},R)=k(\hat{R}-R)^2/\hat{R}^2$ konnten die Variationskoeffizienten der erklärenden Grössen von R_{lat_p} als Bestimmungsparameter der möglichen Informationsprogramme aufgefasst und für das optimale Informationsprogramm schliesslich die Beziehung (59) ermittelt werden.

In beiden Fällen stellten wir noch Betrachtungen über den Spezialfall an, dass die Höhe der Kosten a-priori festgelegt ist und somit nur noch das Problem der optimalen Verwendung des gegebenen Budgets vorliegt. Bei der zweiten Verlustfunktion konnten wir für eine spezielle Kostenfunktion als Lösung dieses Problems die Relation (61) gewinnen.

LITERATURVERZEICHNIS

[1] BARTLETT, M.S. : An Introduction to Stochastic Processes. Cambridge 1956.

[2] CHETTY, V.K. : Bayesian Analysis of Haavelmo's Models Econometrica, 36(1968), S. 582-602.

[3] COCHRAN, W.G. : Sampling Techniques. New York 1953.

[4] DAVIS, H.T. : The Analysis of Economic Time Series. Cowles Comm. Res. Econ. (Bloomington, Indiana) Monogr. no. 6 (1941).

[5] DIEHL, H. und SPROTT, D.A. : Die Likelihoodfunktion und ihre Verwendung beim statistischen Schluss. Statistische Hefte, 6 (1965), S. 112 - 134.

[6] DIEHL, H. und LOUWES, S.L. : Zur Optimierung von Informationsprogrammen bei statistischen Entscheidungen. Statistische Hefte, 9 (1968), S. 176 - 188.

[7] DREZE, J. : The Bayesian Approach to Simultaneous Equations Estimations. O.N.R. Research Memorandum No. 67, September 1962.

[8] DOOB, J.L. : Stochastic Processes. New York 1953.

[9] FISZ, M. : Wahrscheinlichkeitsrechnung und Mathematische Statistik. Berlin 1958.

[10] KENDALL, M.G. und STUART, A. : The Advanced Theory of Statistics. Bd. I: Distribution Theory. London 1958.
Bd. II: Inference and Relationship. London 1961.
Bd.III: Design and Analysis, and Time-Series. London 1966.

[11] LINDLEY, D.V. : Introduction to Probability and Statistics from a Bayesian Viewpoint.
Teil I : Probability. Cambridge 1965.
Teil II: Inference. Cambridge 1965.

[12] LOUWES, S.L. : Economic Policy and Statistics. Neth. J.Agric. Sci., 14 (1966), 1, S. 58 - 69.

[13] LOUWES, S.L. : Cost Allocation in Agricultural Surveys. Review of the International Statistical Institute 35 (1967), 3, S. 264 - 290.

[14] LOUWES, S.L. und DIEHL, H. : Optimum Statistical Programmes for Forecasting Pigmeat Production. Vortrag beim European Meeting of the Econometric Society in Amsterdam 1968.

[15] LUCE, R.D.und RAIFFA, H. : Games and Decisions. New York 1957.

[16] MENGES, G. : Oekonometrie. Wiesbaden 1961 .

[17] MENGES, G. : Oekonometrische Prognosen. Köln-Opladen 1967.

[18] MENGES, G. : Grundriss der Statistik. Teil 1 : Theorie. Köln-Opladen 1968.

[19] MENGES, G. : Grundmodelle wirtschaftlicher Entscheidungen. Köln-Opladen 1969.

[20] MENGES, G. : Vorausschätzung mit Hilfe ökonometrischer Modelle. Allgemeines Statistisches Archiv, 51 (1967), Heft 1, S. 45 - 54.

[21] MENGES, G. und DIEHL, H. : Time Stability of Structural Parameters. Colston Papers, Vol.XVI : Proceedings of the sixteenth Symposium of the Colston Research Society held in the University of Bristol, April 6th - 9th, 1964. London 1964, S. 299 - 317.

[22] MENGES, G. und DIEHL, H. : Das Stabilitätsproblem in der Oekonometrie. Statistische Hefte 6 (1965), S. 27 - 52.

[23] OSTROWSKI, A. : Vorlesungen über Differential- und Integralrechnung.
Band 1: Funktionen einer Variablen. Basel 1952.
Band 2: Differentialrechnung auf dem Gebiete mehrerer Variablen. Basel 1951.
Band 3: Integralrechnung auf dem Gebiete mehrerer Variablen. Basel-Stuttgart 1954.

[24] PRATT, J.W., RAIFFA, H. und SCHLAIFER, R. : Introduction to Statistical Decision Theory. New York - St. Louis - San Francisco - Toronto - London - Sydney 1965.

[25] RAIFFA, H. und SCHLAIFER, R. : Applied Statistical Decision Theory. Boston 1961.

[26] SAVAGE, L.J. : The Foundations of Statistics. New York - London 1954.

[27] SCHNEEWEISS, H. : Das Grundmodell der Entscheidungstheorie. Statistische Hefte 7 (1966), S. 125 - 137.

[28] SCHNEEWEISS, H. : Entscheidungskriterien bei Risiko. Berlin - Heidelberg-New York 1967.

[29] SPROTT, D.A. : Statistical Estimation - some Approaches and Controversies. Statistische Hefte 6 (1965), S. 97 - 111.

[30] STATISTISCHES BUNDESAMT : Stichproben in der amtlichen Statistik. Stuttgart-Mainz 1960.

[31] THEIL, H. : Applied Economic Forecasting. Amsterdam 1966.

[32] WALD, A. : Statistical Decision Functions. New York - London 1950.

[33] WHITTLE, P. : Prediction and Regulation. London 1963.

[34] ZARKOVICH, S.S. : Quality of Statistical Data. FAO Rome 1966.

AMT FÜR AMTLICHE VERÖFFENTLICHUNGEN DER EUROPÄISCHEN GEMEINSCHAFTEN

16753